9 February 2007 | $10

Science

Sustainability and Energy

AAAS

Volume 315, Issue 5813

COVER

A solar farm collects a tiny fraction of the 174,000 terawatts of energy striking Earth in the form of sunlight. In 2005, the rate of global energy consumption was about 16.3 terawatts. See the special section on sustainability and energy beginning on page 781.

Photo illustration: Kelly Krause/Science (Images: Getty/Digital Vision)

>> Editorial p. 737; Science Careers section p. 868
For related online material and Podcast, see p. 731 or go to www.sciencemag.org/sciext/sustainability/

DEPARTMENTS

EDITORIAL

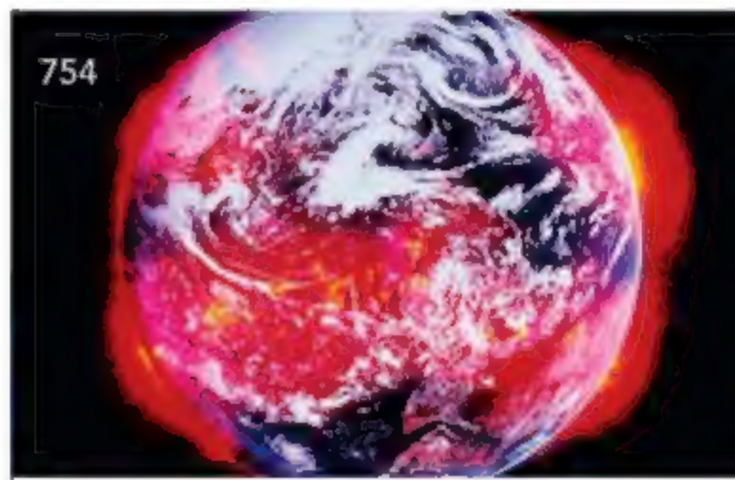
SPECIAL SECTION

Sustainability and Energy

INTRODUCTION

NEWS

PERSPECTIVES

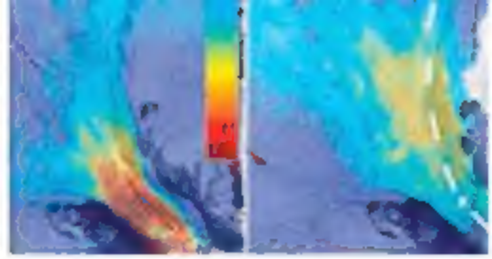


NEWS OF THE WEEK

NEWS FOCUS

CONTENTS continued >>

Science



SCIENCE EXPRESS

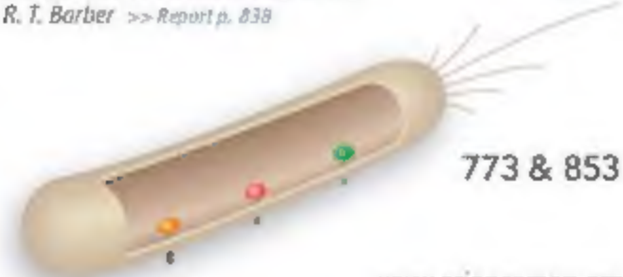
www.sciencexpress.org

LETTERS

BOOKS *ET AL.*

POLICY FORUM

PERSPECTIVES



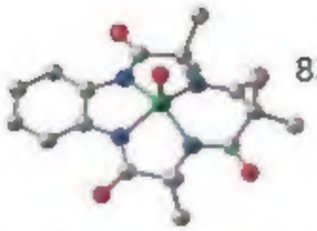
773 & 853

TECHNICAL COMMENT ABSTRACTS

BREVIA

RESEARCH ARTICLES

CONTENTS continued >>



REPORTS

AAAS
ADVANCING SCIENCE, SERVING SOCIETY

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2007 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (53 issues): $142 ($74 allocated to subscription). Domestic institutional subscription (55 issues): $710. Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. **Authorization to photocopy** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $30.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for *Science* is 0036-8075. *Science* is indexed in the Reader's *Guide to Periodical Literature* and in several specialized indexes.

CONTENTS continued >>

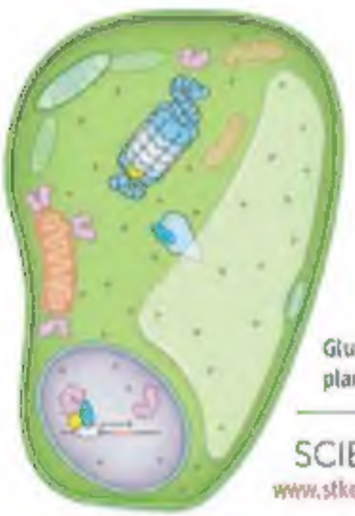
SCIENCENOW

www.sciencenow.org DAILY NEWS COVERAGE

Polar Science From the Deep—Part II
The king crab returns to conquer Antarctica.

Domestic Violence in the Jungle
Male chimps beat females to keep them faithful.

Cracking Concrete's Code
Building material has a surprisingly futuristic makeup.



Glucose regulation of plant gene expression.

SCIENCE'S STKE

www.stke.org

SIGNAL TRANSDUCTION KNOWLEDGE ENVIRONMENT

PERSPECTIVE: Sweet Sensor, Surprising Partners
J.-G. Chen
The glycolytic enzyme hexokinase 1 forms a nuclear complex to regulate the transcription of target genes.

E-LETTER: Breaking T Cell Activation by IKKβ
D. Krappmann
There is more than one way to inhibit Bcl10 function and limit NF-κB activity.

Opportunities in sustainable energy.

SPECIAL SECTION

Sustainability and Energy

SCIENCE CAREERS

www.sciencecareers.org CAREER RESOURCES FOR SCIENTISTS

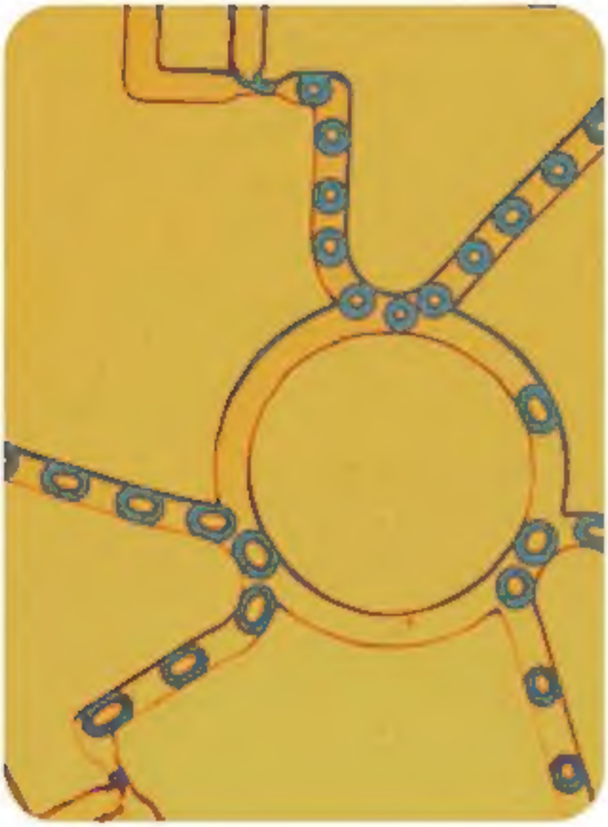
US: From Greener Production to Carbon Trading—Sustainable Energy Careers
J. Austin
Science Careers profiles three scientists doing sustainable energy work in the private sector.

>> Science Careers section p. 868

SCIENCEPODCAST

Listen to the 9 February *Science* Podcast to hear about research and funding challenges for sustainable energy, a European perspective on energy policy, and recent progress in photocatalysis.
www.sciencemag.org/about/podcast.dtl

Separate individual or institutional subscriptions to these products may be required for full-text access.



<< Going Nonlinearly with the Flow

Nonlinear processes can be highly disruptive or turbulent, and thus would be unlikely candidates for being time reversible (see the Perspective by **Epstein**). **Fuerstman *et al.*** (p. 828) exploit the linear and smooth flow associated with microfluidics to create a system that shows nonlinear behavior in the motions of entrained droplets that can be reversed with time. As the fluid flows toward a T-junction, the entrained droplets must choose one branch or the other, and this decision has a nonlinear dependence on the fluid flow and the rate of droplet production. After the two streams recombine, the droplets take on a particular repeat pattern or encoding that can be decoded through a reversal of the fluid flow. **Prakash and Gershenfeld** (p. 832) exploit nonlinear behavior of bubbles in a microfluidic system to create bubble logic, in which the bubbles represent bits of information similar to the ones and zeros used in digital computation. They create a series of simple devices including AND, OR, and NOT gates, as well as more complex signal amplification and processing.

All the Same

The predictable brightnesses of type Ia supernovae make them some of the most useful probes of cosmological distances. Observations of these supernovae suggested the existence of dark energy, but uncertainties linger, and better knowledge of the physics of these stellar explosions is needed to improve their use as distance indicators. By mapping the expansion of the supernova explosion in different spectral lines, **Mazzali *et al.*** (p. 825) show that all type Ia supernovae form from the explosion of a star of the same mass, and their ejecta reach out to similar distances. Simulations of the explosions and analysis of the recorded changes in brightness over time support this finding. These results point to a single explosion mechanism for supernova Type Ia progenitors, supporting their cosmological use.

Small Sinking Surprise

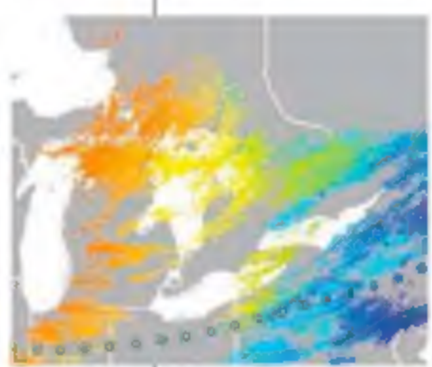
The concentration of CO_2 in the atmosphere is kept in check partly by how much CO_2 is consumed by marine photosynthetic organisms, although the only ones that sequester CO_2 "permanently" are those that sink to the bottom of the ocean and wind up in sediments. It is generally thought that the most important class of organisms responsible for exporting carbon from the surface to deeper waters are the large phytoplankton such as diatoms that can sink rapidly enough to avoid being consumed by microbes on the way down. **Richardson and Jackson** (p. 838; see the Perspective by **Barber**) challenge this assumption by describing how the very small autotrophs, called "picoplankton," that dominate ocean primary production over large regions, can also contribute to export through the processes of aggregation and incorporation. This finding implies that the amount of picoplankton that contribute to the flux of carbon to the deep ocean has been significantly underestimated, and that the importance of these tiny photosynthesizers as "carbon pumps" has been overlooked.

Viral Prefusion Conformation Revealed

The rhabdoviruses, which include human pathogens like rabies virus, enter the cell through the endocytic pathway. Viral membrane fusion with cellular endosomal membranes is triggered by a pH-dependent structural change of a transmembrane viral glycoprotein (G). Now **Roche *et al.*** (p. 843) have determined the prefusion form of the fusion G protein from vesicular stomatitis virus. Comparison with the postfusion structure reveals the structural reorganization between these forms and suggests a pathway for the dramatic, but reversible change.

Directly Observed

Atmospheric deep convection, which transports air from the surface to the upper troposphere is difficult to measure, and our understanding of this process has mainly been based on modeling. **Bertram *et al.*** (p. 816; published online 5 January; see the Perspective by **Jaegle**) provide direct observational constraints to the process, with measurements of a suite of trace gases and aerosols made in the summertime continental upper troposphere over the eastern United States and Canada. Using the distribution of chemical species whose kinetics are well understood to determine the amount of time that air spends in the upper troposphere, they calculated important dynamical parameters such as the extent to which convection perturbs the continental upper troposphere during summer, the fraction of boundary layer air present in convective outflow, and the convective overturn rate of the upper troposphere. These results present a challenge to current ideas about processes that control upper tropospheric ozone and its impact on climate.



Continued on page 735

CREDITS (TOP TO BOTTOM): MANU PRAKASH, FELICE FRANKEL; BERTRAM ET AL.

Continued from page 733

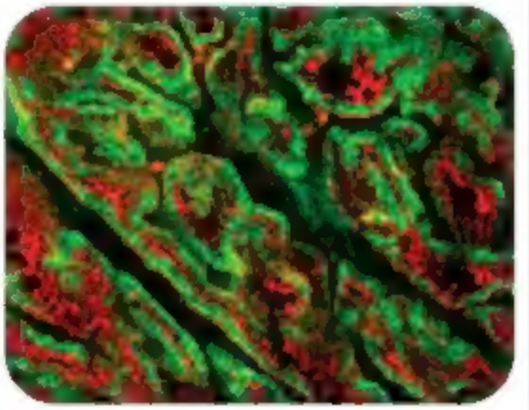
Documenting Diversity

Human genetic diversity includes polymorphisms at individual points within genes, as well as variations in copy number. **Stranger *et al.*** (p. 848) present a whole-genome survey of the effect of such differences on gene expression of nearly 15,000 transcripts. In order to understand complex patterns of inheritance (such as quantitative traits), it will be necessary to look at both types of variation.

Make Your Move

Motility in bacteria is important for pathogenesis and for bacterial chemotaxis. So-called gliding motility in the bacterium *Myxococcus xanthus* is powered by two distinct engines: S motility (powered by type IV pili) and A motility, of unknown mechanism. By following the localization pattern of an A motility–specific protein, **Mignot *et al.*** (p. 853; see the Perspective by **Kearns**) have now discovered that the most popular hypothesis for A motility, directed slime secretion, is likely to be incorrect and that intracellular motor-coupled adhesion complexes power movement. Thus, bacteria and eukaryotic cells may use a similar transient adhesion-based mechanism for motility.

A Receptor for Vitamin A Uptake



Vitamin A has a large number of biological functions, including roles in vision, reproduction, immunity, tissue regeneration, and neuronal signaling. The existence of a cell surface receptor for the vitamin A carrier retinol binding protein (RBP) was proposed more than 30 years ago. **Kawaguchi *et al.*** (p. 820, published online 25 January) now report the identification and characterization of the long-sought RBP receptor. The RBP receptor is a multipass transmembrane domain protein with robust RBP binding and vitamin A uptake activity and is localized to the expected cellular locations for vitamin A uptake.

The Grim Reaper on Autopilot?

The proteins Bax and Bak are key mediators of cell death signals that function at the mitochondria to promote apoptosis. There is evidence for multiple modes of regulation of Bax and Bak in cells. Some studies have proposed that other members of the Bcl-2 family of proteins interact with and directly activate Bax and Bak—a scenario in which cell survival seems to be the default state of the cell. However, activity of Bax and Bak is also held in check by interaction with pro-survival proteins, and it may be that relief of this inhibition determines the cell's fate. **Willis *et al.*** (p. 856; see the Perspective by **Youle**) present evidence that the default state of Bax and Bak would lead to cell death. Cells lacking the direct activators of Bax and Bak still undergo apoptosis in response to overexpression of upstream components of the cell death pathway.

Two Wings Good, Four Wings Better

Some flying insects have two wings, whereas others have four. The common housefly, which possesses two wings, makes use of the vestigial hindwing (the pendulum-shaped haltere) as a source of mechanosensory input to the neural centers that support stable flight. **Sane *et al.*** (p. 863; see the Perspective by **Alexander**) have asked whether moths, which have four wings, possess a similar kind of flight control mechanism. Instead, the antennae appear to serve a haltere-like function by providing mechanosensory input through hairs or bristles located at their base, whose deflections are translated into afferent neural signals.

CREDIT: KAWAGUCHI ET AL.

AAAS*Travels*

Come explore the world with AAAS this year. You will discover excellent itineraries and leaders, and congenial groups of like-minded travelers who share a love of learning and discovery.

Tibetan Plateau
July 4-22, 2007
Explore Tibet, a place of fascination for naturalists and explorers for centuries, from the eastern grasslands to the heart of Tibet—Lhasa & more!

Galapagos Islands
July 21-30, 2007
Discover Darwin's "enchanted isles" on board the *M/V Islander* while exploring the fascinating Galapagos archipelago where wildlife abounds! From $4,150 + air.

A Walk in the Swiss Alps
July 21–August 2, 2007
Discover some of the finest areas in Switzerland for walking: Appenzell and Engelberg, plus see the high alps, Lucerne & St Gallen.

Madagascar
July 24–August 6, 2007
An outstanding introduction to the nature reserves and unique wildlife including lemurs, sifakas, and flying foxes. Visit Perinet, Asole, and Berenty!

Peru & Machu Picchu
July 29–August 8, 2007
Discover the Inca civilization and Peru's cultural heritage with expert **Dr. Douglas Sharon**. Explore Lima, Cuzco, Machu Picchu, the Nazca Lines & more! $3,695 + air.

Xinjiang & Hunza
August 5-22, 2007
Discover the Silk Road in far western China with trip leader **Dr. Chris Carpenter**. Visit the ancient cities of Turpan and Urumqi, legendary Kashgar, and see the Karakoram and Hunza. $3,895 + air.

Call for trip brochures & the Expedition Calendar
(800) 252-4910

AAAS*Travels*
17050 Montebello Road
Cupertino, California 95014
Email: AASinfo@betchartexpeditions.com
On the Web: www.betchartexpeditions.com

John P. Holdren is the president of the AAAS, the director of the Woods Hole Research Center, the Teresa and John Heinz Professor of Environmental Policy at Harvard, and the director of the Program on Science, Technology, and Public Policy in Harvard's Kennedy School of Government.

Energy and Sustainability

THE THEME FOR THE 2007 ANNUAL MEETING OF THE AAAS, TO BE HELD IN SAN FRANCISCO on 15 to 19 February, is "Science and Technology for Sustainable Well-Being." The problem of energy, which is the focus of this special issue of *Science* linked to the meeting, looms as a central element in the web of interwined challenges that this theme embodies.

Well-being has environmental, sociopolitical, and cultural dimensions as well as economic ones, and the goal of sustainable well-being entails improving all of these dimensions in ways and to end points that are consistent with maintaining the improvements indefinitely. This challenge includes not only improving sustainably the standard of living in developing countries, but also converting to a sustainable basis the currently unsustainable practices supporting the standard of living in industrialized ones.

Civilization's ability to meet this immense challenge clearly depends on our strengths in natural science and engineering. But it also depends on our strengths in the social sciences and in "social technology" in the form of business, government, and law, as well as on the societal wit and will to integrate all of these elements in pursuit of the sustainable–well-being goal.

No part of this challenge is more complex or more demanding than its energy dimension. This is so in part because energy supply is tightly intertwined with national and international security and with many of the most damaging and dangerous environmental problems—from indoor air quality to global climate change—as well as with the capacity to meet basic human needs and fuel economic growth.

The multiplicity and importance of these linkages would make energy a vexing issue even in a world where energy demand was constant. But that is not the world we live in. Continuing population growth and rapidly rising affluence in many parts of the globe are driving a rate of increase in energy use that has staggering implications. Even if the energy efficiency of the world economy—gross world product per unit of energy—were to continue to increase at the long-term historical rate of about 1% per year, the realization of middle-of-the-road population and economic projections would entail quadrupling world energy use in this century.

In a world where today one-third of primary energy comes from oil (two-thirds of the remaining high-quality supplies of which probably lie under the volatile Middle East) and 80% comes from oil, coal, and natural gas combined (virtually all of the carbon dioxide from the combustion of which continues to go straight into the atmosphere), that middle-of-the-road energy trajectory cannot be managed simply by expanding what we are already doing. Such a path is not merely unsustainable; it is a prescription for disaster.

The perils of oil dependence and climate change, coupled with the demand for large increases in the per-capita availability of energy services, compel an early transition to a different path. Its requirements include a reduction in global population growth (achievable, fortunately, by means that are desirable in their own right) and a sharply increased emphasis on improving the efficiency of energy conversion and end use (aiming to improve the energy efficiency of the world economy not by 1% per year but by 2% per year or more).

Also required is a severalfold increase in public and private investments to improve the technologies of energy supply. We need to know whether and how the carbon dioxide from fossil-fuel use can be affordably and reliably sequestered away from the atmosphere; whether and how nuclear energy can be made safe enough and proliferation-resistant enough to be substantially expanded worldwide; and to what extent biofuel production can be increased without intolerable impacts on food supply or ecosystem services. And we need to improve the affordability of the direct harnessing of sunlight for society's energy needs.

Much insight about the current prospects in these and other dimensions of the energy problem is available in this special issue of *Science*. Still more about energy is on the agenda for the San Francisco meeting, along with much else germane to "Science and Technology for Sustainable Well-Being." I commend it all to your attention. Nothing is more important to the human condition in the 21st century than rising to this set of challenges.

– John P. Holdren

10.1126/science.1139792

CREDIT: JUPITER IMAGES

EDITED BY GILBERT CHIN AND JAKE YESTON

GEOLOGY

Leveling the Landscape

It has been recognized for some time that humans are the most important geomorphic agent modifying our planet's surface, dwarfing the effects of deforestation, desertification, and erosion caused by other processes globally. However, the details of these activities are enlightening, and Wilkinson and McElroy compare the rates of erosion from human activity in different settings with natural processes and with long-term and short-term rates inferred throughout Earth's history. About 5 gigatons (Gt) of sediment per year are thought to have been deposited naturally by rivers during the past 540 million years. There has been high variation about this average, particularly since the Pliocene and during glacial times; the average flux today is about 21 Gt/year. Most of this material (about 80%) comes from mountainous regions where natural erosion rates are highest. For comparison, it is estimated that humans now move about 75 Gt of dirt and rock annually, mostly in low-lying or low-topography areas, and particularly near coasts. This difference implies that large amounts of human-derived detritus are being stored primarily on floodplains and in small stream networks on coastal plains immediately downslope from such areas. This flux greatly exceeds the movement of material by Pleistocene ice sheets. — BH

Geol. Soc. Am. Bull. **119**, 140 (2007).

MOLECULAR BIOLOGY

Remodeling with a Floral Motif

The transition from vegetative growth to flowering in the plant *Arabidopsis* is in part regulated by the gene *Flowering Locus C* (*FLC*). The expression of *FLC* is known to be regulated by various signaling pathways that modify the chromatin in which it is embedded. Deal *et al.* show that, beyond secondary modifications of existing chromatin components, the addition of a specific histone variant, H2A.Z, adds to the controls on *FLC* expression. Histone H2A.Z is deposited onto *FLC* chromatin by the *Arabidopsis* variant of the SWR1 complex, which has been shown in yeast to participate in the remodeling and individualizing of local chromatin domains. With histone H2A.Z as part of its chromatin packaging, the *FLC* gene is better poised for active expression. It is the robust expression of *FLC* that results in maintenance of the vegetative state. — PJH

Plant Cell **19**, 10.1105/tpc.106.048447 (2007).

ECOLOGY/EVOLUTION

Managing Murrelets

One challenge in conservation management is estimating what a sustainable population should look like. Geographic and genetic information can readily be obtained from museum specimens, but if, as for many birds, there are morphologically distinguishable age classes, then age-ratio analysis can also provide baseline rates for reproduction and survival. The output of such an analysis can be used to set targets for population recovery.

Beissinger and Peery have championed the case of the marbled murrelet (*Brachyramphus marmoratus*), an endangered seabird from the Californian coast which unexpectedly nests high up in old-growth coniferous trees and whose plight has only recently come to attention. The murrelet population is being decimated as a result of attacks from crows, by logging, and especially via the loss of the fish that they eat. Reproduction is expensive for murrelets; they lay a single egg that weighs up to a quarter of the adult bird's weight, and adults will abandon breeding in the face of insufficient food. There are some methodological risks with age-ratio analysis, but for this bird, contemporary data from field studies were compared with the museum data, which showed that the reproductive success of contemporary murrelets is almost an order of magnitude less than it was a century ago. — CA

Marbled murrelet.

Ecology **88**, 296 (2007).

CREDITS (TOP TO BOTTOM): JIM ZUCKERMAN/CORBIS; BEISSINGER AND PEERY, *ECOLOGY* **88**, 296 (2007)

CHEMISTRY

Bromine Scrub

One drawback of power generation from coal combustion is the atmospheric release of toxic mercury previously trapped in the solid coal. Although current emission filtration systems efficiently remove oxidized mercury compounds, they are less effective at trapping the metal in its water-insoluble elemental state. Combustion of chlorine-rich bituminous coals tends to pass less mercury through the filters, an observation that has been attributed to mercury oxidation by the chlorine. Liu *et al.* therefore explored the capacity of bromine, chlorine's more polarizable congener, to oxidize residual Hg(0) in coal emission streams. Laboratory-scale tests revealed a greater than hundredfold enhancement in the oxidation rate relative to measured chlorination kinetics. Moreover, the reaction was further accelerated by adsorption of the reagents onto fly-ash particles, with little or no inhibition observed from oxygen, CO, water, or HCl. Sulfur dioxide proved mildly inhibitory, whereas NO had a small promoting effect on the fly-ash–mediated oxidation. By extrapolating their data to the higher temperatures of a power plant emission

Continued on page 741

Continued from page 739

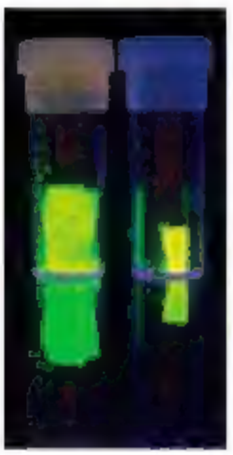
stream, the authors estimate that the addition of 0.4 part per million of bromine to flue gas could oxidize 60% of residual Hg(0). — JSY

Environ. Sci. Technol. **41**, 10.1021/es061705p (2007).

CHEMISTRY

All Glow Together

The appeal of quantum dots stems from their highly tunable optical and electronic properties. Their size-dependent luminescence, for example, is finding use in biological sensing and imaging. However, with the exception of this application, most technologies that could take advantage of quantum dots do not operate in solution. It can be difficult to pack a large concentration of quantum dots into a solid structure, and even if assembly is successful, the luminescent properties can be lost because of the formation of trap states or to quenching interactions between neighboring dots. Arachchige and Brock used a sol-gel process to prepare aerogel and xerogel monoliths with CdSe/ZnS core-shell nanoparticles as the primary building blocks. Throughout the transformation process from sols to wet gels and on to monoliths, little change was observed in the photoluminescence spectra for the three sizes of particles tested. This behavior contrasts strongly with that of nanostructured networks assembled from bare CdSe particles, which lost the favorable absorption and emission properties observed in solution. In both the aerogels and xerogels, transmission electron microscopy images suggested that the ZnS shells may have partially fused together to form a matrix containing CdSe particles, particularly in the 20-fold denser xerogels. — MSL



Luminescent aerogel (left) and xerogel (right).

CREDITS: ARACHCHIGE AND BROCK, *J. AM. CHEM. SOC.* **129**, 10.1021/JA066749C (2007)

J. Am. Chem. Soc. **129**, 10.1021/ja066749c (2007).

BIOMEDICINE

Viruses in the Prostate

Twenty to 25% of human cancers are caused by viruses. The recognition that a virus plays an etiological role in a specific type cancer can profoundly change the ways in which that particular disease is diagnosed, treated, or prevented. One recent guidepost is cervical cancer, where the identification of human papillomavirus as a causal agent has led to the development of a promising prophylactic vaccine.

Previous studies have provided evidence in support of a possible viral origin for human prostate cancer. A retrovirus called XMRV (for xenotropic murine leukemia virus–related virus) was detected in 40% of prostate tumors from men who were homozygous for an allelic variant of the *RNASEL* gene and in only 2% of tumors from men of other genotypes. The *RNASEL* gene encodes RNase L, a ribonuclease whose activity is required for an innate antiviral response mediated by interferon (IFN); and, intriguingly, the allelic variant associated with XMRV-containing tumors encodes an enzyme with impaired activity.

Dong *et al.* show that a molecular viral clone of XMRV is infectious in human prostate cancer cell lines, that replication of the virus in vitro is sensitive to inhibition by IFN, and that suppression of RNase L enhances viral replication. In addition, they localized putative integration sites for the XMRV provirus to several host genes that encode functions with biologically plausible roles in prostate cancer, including a suppressor of androgen receptor transactivation. Still unanswered is the critical question of whether XMRV plays a causal role in prostate cancer, but these provocative observations should stimulate further experiments to sort this out. — PAK

Proc. Natl. Acad. Sci. U.S.A. **104**, 1655 (2007).

GEOPHYSICS

Rough Core

Earth's core is split into two zones. A solid inner core lies within a liquid outer core, which drives the geodynamo via heat convection. The smooth and spherical interface comprising the inner core boundary (ICB) is the site of such geochemical changes as solidification and chemical flows. The ICB can be probed by seismic waves that reflect or refract off of it. In the unusual circumstance of repeating earthquakes, sequential rays can travel identical ray paths through Earth. By analyzing one such pair of repeated earthquakes, Cao *et al.* found changes in the ICB structure over a decade. Two identical ray paths, sampled by earthquakes in the South Sandwich Islands in 1993 and 2003 and recorded in Alaska and Canada, show variations in the shape of their seismic response over this time. Although simple pressure waves were identical, differences were seen between the reflected and refracted waves that intercepted the ICB. These amplitude differences may indicate lumps on the ICB of 10 km in horizontal extent. Features of this size could be created by low convection in the inner core itself, or by rotation of a fixed bump on the inner core by approximately 0.1° per year. — JB

Proc. Natl. Acad. Sci. U.S.A. **104**, 31 (2007).

Institutional Site License Available

What can *Science* STKE give me?

A

The definitive resource on cellular regulation

STKE – Signal Transduction Knowledge Environment offers:

- A weekly electronic journal
- Information management tools
- A lab manual to help you organize your research
- An interactive database of signaling pathways

STKE gives you essential tools to power your understanding of cell signaling. It is also a vibrant virtual community, where researchers from around the world come together to exchange information and ideas. For more information go to www.stke.org

To sign up today, visit promo.aaas.org/stkeas

Sitewide access is available for institutions. To find out more e-mail stkelicense@aaas.org

Science stke AAAS

www.sciencemag.org

Science

1200 New York Avenue, NW
Washington, DC 20005
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6500, FAX 202-371-9227

Bateman House, 82-88 Hills Road
Cambridge, UK CB2 1LQ
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

Subscription Services For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

Institutional Site Licenses please call 202-326-6755 for any questions or information

Reprints: Author Inquiries 800-635-7181
Commercial Inquiries 803-359-4578
Corrections 202-326-6501

Permissions 202-326-7074, FAX 202-682-0816

Member Benefits Bookstore: AAAS/BarnesandNoble.com bookstore www.aaas.org/bn; Car purchase discount: Subaru VIP Program 202-326-6417; Credit Card: MBNA 800-847-7378; Car Rentals: [illegible]; AAAS Travels: Betchart Expeditions 800-252-4910; Life Insurance: Seabury & Smith 800-424-9883; Other Benefits: AAAS Member Services 202-326-6417 or www.aasmember.org.

science_editors@aaas.org	(for general editorial queries)
science_letters@aaas.org	(for queries about letters)
science_reviews@aaas.org	(for returning manuscript reviews)
science_bookrevs@aaas.org	(for book review queries)

Published by the American Association for the Advancement of Science (AAAS), *Science* serves as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by the AAAS or the institutions with which the authors are affiliated.

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science and innovation throughout the world for the benefit of all people. The goals of the association are to: foster communication among scientists, engineers and the public; enhance international cooperation in science and its applications; promote the responsible conduct and use of science and technology; foster education in science and technology for everyone; enhance the science and technology workforce and infrastructure; increase public understanding and appreciation of science and technology; and strengthen support for the science and technology enterprise.

Information for Authors

See pages 120 and 121 of the 5 January 2007 issue or access www.sciencemag.org/feature/contribinfo/home.shtml

Editor-in-Chief **Donald Kennedy**
Executive Editor **Monica M. Bradford**
Deputy Editors **R. Brooks Hanson, Barbara R. Jasny, Katrina L. Kelner**
News Editor **Colin Norman**

Editorial Supervisory Senior Editor Phillip D. Szuromi; Senior Editor/Perspectives Lisa D. Chong; Senior Editors [illegible] Pamela J. Hines, Paula A. Kiberstis (Boston), [illegible], Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Valda Vinson, David Voss; Associate Editors Jake S. Yeston, Laura M. Zahn; Online Editor Stewart Wills; Associate Online Editor Tara S. Marathe; Book Review Editor Sherman J. Suter; Associate Letters Editor Etta Kavanagh; Information Specialist Janet Kegg; Editorial Manager Cara Tate; Senior Copy Editors Jeffrey E. Cook, Cynthia Howe, Harry Jach, Barbara P. Ordway, [illegible]; Copy Editors [illegible]; Editorial Coordinators [illegible] Beverly Shields; Publication Assistants [illegible]; Editorial Assistants [illegible]; Executive Assistant [illegible]; Administrative Support [illegible]

News Senior Correspondent [illegible]; Deputy News Editors Robert [illegible]; Contributing Editors [illegible]; News Writers [illegible]; Contributing Correspondents Barry A. Cipra, Jon Cohen (San Diego, CA), Daniel Ferber, Ann Gibbons, Robert Irion, Mitch Leslie, Charles C. Mann, Evelyn Strauss, Gary Taubes, Ingrid Wickelgren; Copy Editors Linda B. Felaco, Rachel Curran, Sean Richardson; Administrative Support Scherraine Mack, Fannie Groom; Bureaus: Berkeley, CA: 510-652-0302, FAX 510-652-1867, New England: 207-549-7755, San Diego, CA: 760-942-3252, FAX 760-942-4979, Pacific Northwest: 503-963-1940

Production Director James Landry; Senior Manager Wendy K. Shank; Assistant Manager Rebecca Doshi; Senior Specialists Jay Covert, Chris Redwood; [illegible]; Preflight Director David M. Tompkins; Manager Marcus Spiegler; Specialist Jessie Mudjitaba

Art Director [illegible]; Associate Art Director Aaron Morales; Illustrators [illegible]; Art Associates [illegible]; Associate Photo Editor [illegible]; Photo Editor Leslie Blizard

Science International

Europe [illegible]; Editorial Supervisory Senior Editor Andrew M. Sugden; Senior Editor/Perspectives Julia Fahrenkamp-Uppenbrink; Senior Editors Caroline Ash (Geneva: +41 (0) 222 346 3106), Stella M. Hurtley, Ian S. Osborne, Stephen J. Simpson, Peter Stern; Associate Editor Joanne Baker; Editorial Support Alice Whaley, Deborah Dennison; Administrative Support Janet Clements, Phil Marlow, H. White; News: Europe News Editor Daniel Clery; Correspondent Gretchen Vogel (Berlin: +49 (0) 30 2809 3902, FAX +49 (0) 30 2809 8365); Contributing Correspondents Michael Balter (Paris), Martin Enserink (Amsterdam and Paris), [illegible]

Asia Japan Office: [illegible]; Asia News Editor Richard Stone [illegible]; Japan News Bureau: Dennis Normile (contributing correspondent, +81 (0) 3 3391 0630, FAX 81 (0) 3 5936 3531; dnormile@gol.com); China Representative Hao Xin, + 86 (0) 10 6307 4439 or 6307 3676, FAX +86 (0) 10 [illegible]; South Asia Pallava Bagla [illegible]

Africa Robert Koenig (contributing correspondent, rob.koenig@gmail.com)

Executive Publisher **Alan I. Leshner**
Publisher **Beth Rosner**

Fulfillment & Membership Services (membership@aaas.org) Director Marlene Zendell; Manager Waylon Butler; Systems Specialist Andrew Vargo; Customer Service Supervisor Pat Butler; Specialists Laurie Baker, Tamara Alfson, Karena Smith, Vicki Linton, Latoya Casteel; Circulation Associate [illegible]; Data Entry Supervisor Cynthia Johnson; Specialists Tomeka Diggs, Tarrika Hill, Erin Layne

Business Operations and Administration Director Deborah Rivera-Wienhold; Business Manager Randy Yi; Senior Business Analyst Lisa Donovan; Business Analyst Jessica Tierney; Financial Analyst Michael LoBue, Farida Yeasmin; Rights and Permissions: Administrator Emilie David; Associate Elizabeth Sandler; Marketing: Director John Meyers; Marketing Managers Darryl Walter, Allison Pritchard; Marketing Associates Julianne Wielga, Mary Ellen Crowley, Catherine Featherston, Alison Chandler, Lauren Lamoureux; International Marketing Manager Wendy Sturley; Marketing Executive Jennifer Reeves; Marketing/Member Services Executive Linda Rusk; Japan Sales Jason Hannaford; Site License Sales Director Tom Ryan; Sales and Customer Service Mehan Dossani, Kiki Forsythe, Catherine Holland, Wendy Wise; Electronic Media: Manager Lizabeth Harman; Project Manager Trista Snyder; Assistant Manager Lisa Stanford; Production Associates Nichele Johnston, Kimberly Oster

[illegible]

Product (science_advertising@aaas.org); Midwest Rick Bongiovanni: 330-405-7080, FAX 330-405-7081 • West Coast/W. Canada Teola Young: 650-964-2266 • East Coast/E. Canada Christopher Breslin: 443-512-0330, FAX 443-512-0331 • UK/Europe/Asia Julie Skeet: +44 (0) 1223-326-524, FAX +44 (0) 1223-325-532 • Japan Mashy Yoshikawa: +81 (0) 33235 5961, FAX +81 (0) 33235 5852 • Traffic Manager Carol Maddox; Sales Coordinator Dejandra Simms

Commercial Editor Sean Sanders: 202-326-6430

Classified (advertise@sciencecareers.org); U.S.: Recruitment Sales Manager Ian King: 202-326-6528, FAX 202-289-6742; Inside Sales Manager Daryl Bryant: 202-326-6533; [illegible] Daryl Anderson: 202-326-6543; Northeast: Allison Millar: 202-326-6572; Southeast: Fernando Junco: 202-326-6740; West: Katie Putney: 202-326-6577; Sales Coordinators Erika Bryant, Rohan Edmonson, Shirley Young; International: Sales Manager Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532; Sales Christina Harrison, Svitlana Barnes; Sales Assistant [illegible]; Japan: [illegible]; Advertising Production Operations Manager Deborah Tompkins; Associates Christine Hall, Amy Hardcastle; Publications Assistants Robert Buck, Mary Lagnaoui

AAAS Board of Directors Retiring President, Chair Gilbert S. Omenn; President John P. Holdren; President-Elect David Baltimore; Treasurer David E. Shaw; Chief Executive Officer Alan I. Leshner; Board Rosina M. Bierbaum, John E. Dowling, Lynn W. Enquist, Susan M. Fitzpatrick, Alice Gast, Thomas Pollard, Peter J. Stang, Kathryn D. Sullivan

Senior Editorial Board

[illegible], Chair, Stanford Univ.; Richard [illegible]; Robert May, [illegible]; Marcia McNutt, Monterey Bay Aquarium Research Inst.; Linda Partridge, [illegible]; Vera C. Rubin, Carnegie Institution of Washington; Christopher R. Somerville, Carnegie Institution; George M. Whitesides, Harvard University

Board of Reviewing Editors

Joanna Aizenberg, Bell Labs/Lucent; R. McNeill Alexander, Leeds Univ.; David Altshuler, Broad Institute; Arturo Alvarez-Buylla, Univ. of California, San Francisco; Richard Amasino, Univ. of Wisconsin, Madison; Meinrat O. Andreae, Max Planck Inst., Mainz; Kristi S. Anseth, Univ. of Colorado; John A. Bargh, Yale Univ.; Cornelia I. Bargmann, Rockefeller Univ.; Brenda Bass, Univ. of Utah; Marisa Bartolomei, Univ. of Penn. School of Med.; Ray H. Baughman, Univ. of Texas, Dallas; Stephen J. Benkovic, Pennsylvania St. Univ.; Michael J. Bevan, Univ. of Washington; Ton Bisseling, Wageningen Univ.; Mina Bissell, Lawrence Berkeley National Lab; Peer Bork, EMBL; Dianna Bowles, Univ. of York; Robert W. Boyd, Univ. of Rochester; Dennis Bray, Univ. of Cambridge; Stephen Buratowski, Harvard Medical School; Jillian M. Buriak, Univ. of Alberta; Joseph A. Burns, Cornell Univ.; William P. Butz, Population Reference Bureau; Peter Carmeliet, Univ. of Leuven, VIB; Gerbrand Ceder, MIT; Mildred Cho, Stanford Univ.; David Clapham, Children's Hospital, Boston; David Clary, Oxford University; J. M. Claverie, CNRS, Marseille; Jonathan D. Cohen, Princeton Univ.; Stephen M. Cohen, [illegible]; Robert H. Crabtree, Yale Univ.; F. Fleming Crim, Univ. of Wisconsin; William Cumberland, UCLA; George Q. Daley, Children's Hospital, Boston; Edward DeLong, MIT; Emmanouil T. Dermitzakis, Wellcome Trust Sanger Inst.; Robert Desimone, MIT; Dennis Discher, Univ. of Pennsylvania; W. Ford Doolittle, Dalhousie Univ.; Jennifer A. Doudna, Univ. of California, Berkeley; Julian Downward, Cancer Research UK; Denis Duboule, Univ. of Geneva; Christopher Dye, WHO; Richard Ellis, Cal Tech; Gerhard Ertl, Fritz-Haber-Institut, Berlin; Douglas H. Erwin, Smithsonian Institution; Barry Everitt, Univ. of Cambridge; Paul G. Falkowski, Rutgers Univ.; Ernst Fehr, Univ. of Zurich; Tom Fenchel, Univ. of Copenhagen; Alain Fischer, INSERM; Jeffrey S. Flier, Harvard Medical School; Chris D. Frith, Univ. College London; John Gearhart, Johns Hopkins Univ.; Wulfram Gerstner, [illegible] Inst. of Technology; H. C. J. Godfray, Univ. of Oxford; Jennifer M. Graves, Australian National Univ.; Christian Haass, Ludwig Maximilians Univ.; Dennis L. Hartmann, Univ. of Washington; Chris Hawkesworth, Univ. of Bristol; Martin Heimann, Max Planck Inst., Jena; James A. Hendler, Univ. of Maryland; Ray Hilborn, Univ. of Washington; Ove Hoegh-Guldberg, Univ. of Queensland; Ary A. Hoffmann, La Trobe Univ.; Ronald R. Hoy, Cornell Univ.; Evelyn L. Hu, Univ. of California, SB; Olli Ikkala, Helsinki Univ. of Technology; Meyer B. Jackson, Univ. of Wisconsin Med. School; Stephen Jackson, Univ. of Cambridge; Steven Jacobsen, Univ. of California, Los Angeles; Peter Jonas, Universität Freiburg; Daniel Kahne, Harvard Univ.; Bernhard Keimer, Max Planck Inst., Stuttgart; Elizabeth A. Kellog, Univ. of Missouri, St. Louis; Alan B. Krueger, Princeton Univ.; Lee Kump, Penn State; Mitchell A. Lazar, Univ. of Pennsylvania; Virginia Lee, Univ. of Pennsylvania; Anthony J. Leggett, Univ. of Illinois, Urbana-Champaign; Michael J. Lenardo, NIAID, NIH; Norman L. Letvin, Beth Israel Deaconess Medical Center; Olle Lindvall, Univ. Hospital, Lund; Richard Losick, Harvard Univ.; Ke Lu, [illegible]; Andrew P. MacKenzie, Univ. of St. Andrews; Raul Madariaga, École Normale Supérieure, Paris; Anne Magurran, Univ. of St. Andrews; Michael Malim, King's College, London; Virginia Miller, Washington Univ.; Yasushi Miyashita, Univ. of Tokyo; Richard Morris, Univ. of Edinburgh; Edvard Moser, Norwegian Univ. of Science and Technology; Andrew Murray, Harvard Univ.; Naoto Nagaosa, Univ. of Tokyo; James Nelson, Stanford Univ. School of Med.; Roeland Nolte, Univ. of Nijmegen; Helga Nowotny, European Research Advisory Board; Eric N. Olson, Univ. of Texas, SW; Erin O'Shea, Harvard Univ.; Elinor Ostrom, Indiana Univ.; Jonathan T. Overpeck, Univ. of Arizona; John Pendry, Imperial College; Philippe Poulin, CNRS; Mary Power, Univ. of California, Berkeley; Molly Przeworski, Univ. of Chicago; David J. Read, Univ. of Sheffield; Les Real, Emory Univ.; Colin Renfrew, Univ. of Cambridge; Trevor Robbins, Univ. of Cambridge; Barbara A. Romanowicz, Univ. of California, Berkeley; Nancy Ross, Virginia Tech; Edward M. Rubin, Lawrence Berkeley National Lab; J. Roy Sambles, Univ. of Exeter; Jürgen Sandkühler, Medical Univ. of Vienna; David S. Schimel, National Center for Atmospheric Research; Georg Schulz, Albert-Ludwigs-Universität; Paul Schulze-Lefert, Max Planck Inst., Cologne; Terrence J. Sejnowski, The Salk Institute; David Sibley, Washington Univ.; Montgomery Slatkin, Univ. of California, Berkeley; George Somero, Stanford Univ.; Joan Steitz, Yale Univ.; Elsbeth Stern, ETH Zürich; Thomas Stocker, Univ. of Bern; Jerome Strauss, Virginia Commonwealth Univ.; Marc Tatar, Brown Univ.; Glenn Telling, Univ. of Kentucky; Marc Tessier-Lavigne, Genentech; Michiel van der Klis, Astronomical Inst. of Amsterdam; Derek van der Kooy, Univ. of Toronto; Bert Vogelstein, Johns Hopkins; Christopher A. Walsh, Harvard Medical School; Graham Warren, Yale Univ. School of Med.; Colin Watts, Univ. of Dundee; Julia R. Weertman, Northwestern Univ.; Jonathan Weissman, Univ. of California, San Francisco; Ellen D. Williams, Univ. of Maryland; R. Sanders Williams, Duke University; Ian A. Wilson, The Scripps Res. Inst.; Jerry Workman, Stowers Inst. for Medical Research; John R. Yates III, The Scripps Res. Inst.; Martin Zatz, NIMH, NIH; Huda Zoghbi, Baylor College of Medicine; Maria Zuber, MIT

Book Review Board

John Aldrich, Duke Univ.; David Bloom, Harvard Univ.; Angela Creager, Princeton Univ.; Richard Shweder, Univ. of Chicago; Ed Wasserman, DuPont; Lewis Wolpert, Univ. College, London


(Masturbator in a Surrealist Landscape with D.N.A.)

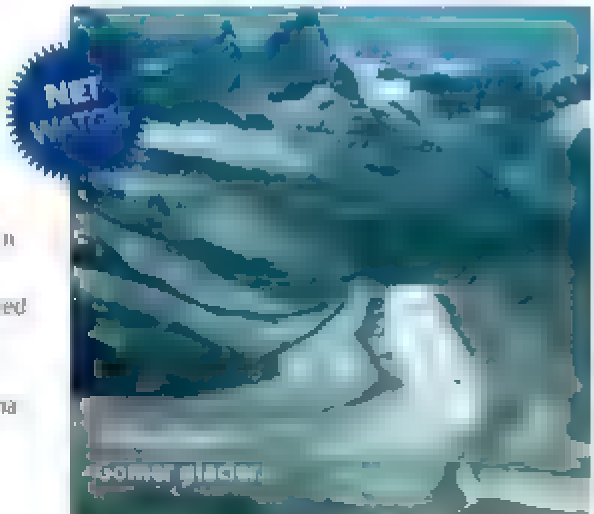
Dalí Postmortem

Eighteen years after the death of Salvador Dalí, some of his genetic material lives on. Michael Rieders, a Pennsylvania forensic toxicologist and longtime fan of the Spanish surrealist, decided to recover the DNA of this painter who "liked to make art out of scientific discoveries." Friends of Dalí's in France sent Rieders two feeding tubes that had kept the artist alive during his last days. After extracting and fingerprinting DNA from both tubes, Rieders found that the samples were identical and free of any other genetic material. So they had to be Dalí's. Rieders plans to send the DNA to galleries and museums to help them authenticate disputed Dalí works in the future.

NIDA Versus Wikipedia

Want to improve your public image? Try rewriting your entry on the Wikipedia Web site. At least that's the tack taken by someone at the U.S. National Institute on Drug Abuse (NIDA). Last year, an anonymous employee of the agency (indicated by the IP address) repeatedly removed controversial sections of an article on NIDA, replacing them with prose about "unprecedented opportunities" at the institute and its aim of "improving the health of the Nation." Among the sections deleted during a heated editing battle with Wikipedians, the citizens who write and monitor the entries, were mentions of debates over the potency of marijuana given to NIDA-funded weed researchers at the University of Mississippi and complaints that government statistics on emergency room visits overstate the dangers of pot.

Deleted sections have since been restored. But NIDA maintains that the write-up is "biased" and "rife with inaccuracies," including outdated information and "information on non-NIDA funded research," and "only touches on a small subset" of the agency's research portfolio.



Gomer glacier

Watching the Glaciers Go

Dwindling mountain glaciers provide one line of evidence for global warming. Where are they shrinking fastest? To find out, slide over to this data storehouse from the World Glacier Monitoring Service (WGMS) in Zurich, Switzerland.

Just posted are the 2004 and 2005 measurements of net change in ice thickness, or mass balance, for 100 glaciers, including South Cascade in Washington state, Saint Sorlin in France, and Bahía del Diablo in Antarctica. Researchers have been continuously monitoring 30 mountain glaciers since 1980. The thickness of the average glacier fell by 725 millimeters in 2004 and another 625 millimeters in 2005. Only 20% of the glaciers are enlarging.

For results on several hundred glaciers going back to 1990, you can download reports, issued every 5 years, that record variables such as movement of a glacier's front. One limitation of the site is that it doesn't provide a complete archive—to get data extending back to 1959, you have to contact WGMS. >> [illegible]


Infrared image of cool top.

<< No Sweat

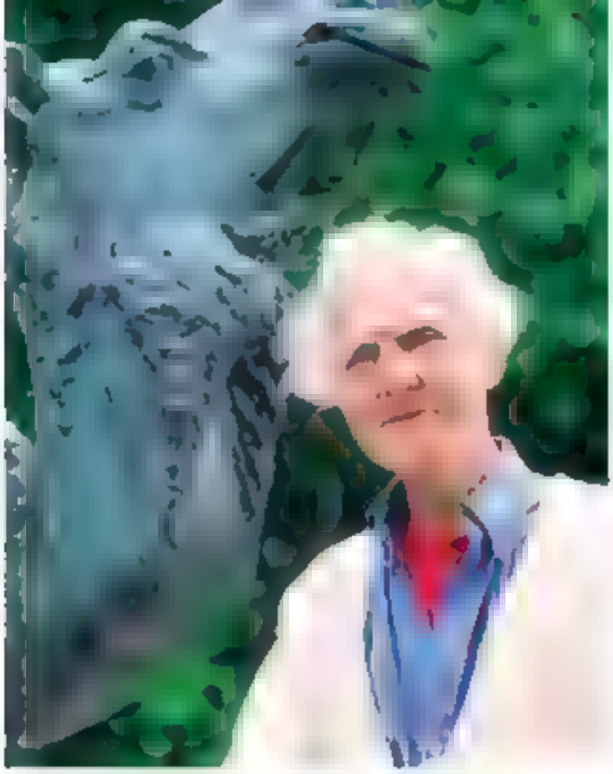
By tradition, a hockey player's jersey is called a sweater, but National Hockey League (NHL) players are getting new togs that will minimize perspiration.

At the request of Reebok, which manufactures uniforms for the NHL, researchers at Central Michigan University in Mount Pleasant tested three different designs against a conventional sweater to see which kept a teenage player coolest during a simulated workout. "The idea is to find the combination of fiber, yarn, and structure that maximizes heat dissipation," says Maureen MacGillivray, a functional apparel designer. She and colleagues used a thermal camera to map the surface temperatures of the jersey, the underlying pads, an undergarment, and the subject's skin. They also used a scanner to make 3D images to study how the tighter fitting sweaters moved with the athlete. They found that the best design reduced skin temperature by as much as 3.3°C. The new material, which is engineered to transport moisture away from the body, helped keep the player cooler, but the cut of the garment was also important, says MacGillivray.

Reebok kept close tabs on its prototypes, MacGillivray says. "People on campus say, 'Oh, we want to see the new jersey,' but they took them all back." NHL players will wear them starting next season. Next job for MacGillivray and colleagues: cooler basketball uniforms.

CREDITS (TOP TO BOTTOM): SALVADOR DALI/GALLERY W. HAEBERLI, SWITZERLAND; T. DOMINA/CMU

EDITED BY VUDHIJIT BHATTACHARJEE



Deaths

FOREVER CURIOUS. The title of one of his books, *How Animals Work*, captures the essence of biologist Knut Schmidt-Nielsen, who died in Durham, North Carolina on 25 January at the age of 91. Considered the father of comparative physiology and integrative biology, the Duke University researcher endured the heat of the Sahara and the cold of the Arctic to learn how animals thrive in extreme climates.

He discovered, for example, that moisture-conserving mucus in the nose—not water stored in its hump—helps protect camels against dehydration, and that special glands help seabirds and marine reptiles shed excess salt. His lab was a menagerie, including a misnamed 100-kilogram female ostrich called Pete who was a source for 1-kilo omelets. (Pete helped Schmidt-Nielsen and his colleagues learn how ostriches can run in the heat without sweating.) "Knut invoked in all of us a sense of curiosity about how animals function," says Barbara Block of Stanford University in Palo Alto, California.

IN BRIEF

■ **Harpal Kumar** is the new CEO of Cancer Research UK, a British nonprofit that provides $500 million a year for cancer research. He succeeds chemist Alex Markham.

■ The **Prince of Wales** has won the Global Environmental Citizen Award from Harvard Medical School for his efforts to promote sustainable farming and environmentally friendly urban development. Prince Charles received the honor last month from actress Meryl Streep and former U.S. vice president Al Gore in New York City.

■ **Jeannette Wing** of Carnegie Mellon University in Pittsburgh, Pennsylvania, has been named head of the National Science Foundation's $527 million Computer and Information Science and Engineering directorate. She replaces Peter Freeman, who has joined a Washington, D.C., consulting firm.

MOVERS

TRADING CONTINENTS. Jonathan Slack, a British developmental biologist, is coming to the University of Minnesota next month to replace Catherine Verfaillie as head of its Stem Cell Institute. The Minnesota institute has close ties with the Stem Cell Institute at the University of Leuven, Belgium, which Verfaillie now heads.

Slack, 58, chairs the department of biology and biochemistry at the University of Bath, where he established a Centre for Regenerative Medicine. Describing stem cell research as "the applied science of developmental biology," Slack says he plans to continue his work both on tail regeneration in the *Xenopus* tadpole and on reprogramming adult cells, such as liver cells, into pancreatic islet cells to treat diabetes.

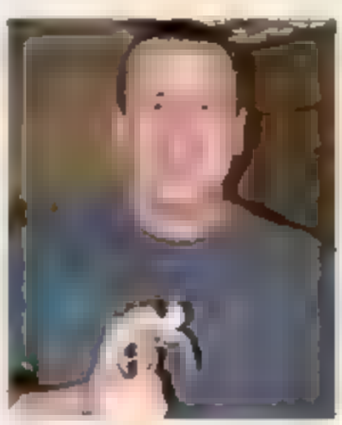
AWARDS

BURNING BRIGHT. Brian Warner had always wanted to be an astronomer. And although circumstances led him into jobs in radio and TV as a news director instead, Warner still managed to spend countless late nights tracking the waxing and waning of asteroid brightness from his backyard observatory north of Colorado Springs, Colorado. Last week, his contributions earned him the American Astronomical Society's first-ever Chambliss Amateur Achievement Medal, a silver medal and a $1000 honorarium.

Warner, 54, now a computer programmer, has published more than 200 records of varying asteroid brightness using data captured with cheap but ultrasensitive light detectors. A light curve reveals not only the shape and rotation of asteroids but also whether an apparent solitary asteroid is actually a pair. Warner's discovery of numerous pairs in the main asteroid belt has challenged theorists to explain how binary asteroids could form there.

Slice of Life >>

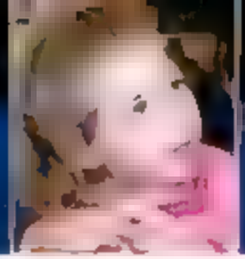
BREAKING THE SILOS. As a postdoc, newly minted Ph.D. physicist Daniel Goldman broke fresh interdisciplinary ground by applying his expertise in complex fluids and granular materials to the study of animal locomotion. But despite the innovativeness of his research—or perhaps because of it—his 6-month search for an academic job posed a challenge. Physics departments said his work wasn't fundamental enough, and biology departments clucked that his work wasn't aimed at basic questions such as how cells work. One interviewer, he says, told him "that I would be working on an island" with few connections on campus.



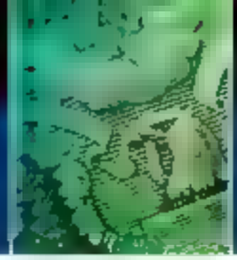
But his perseverance has paid off. Last month, he won a tenure-track slot in the physics department at Georgia Institute of Technology in Atlanta. "We expect that he will build a successful program ... that will make unique contributions to the study of locomotion biology and biomechanics," says Mei-Yin Chou, his new department chair.

Got a tip for this page? E-mail people@aaas.org

CREDITS (TOP TO BOTTOM): DUKE UNIVERSITY NEWS SERVICE; BRIAN WARNER; DANIEL GOLDMAN



Can Rett syndrome be reversed? 749



Science and the U.S. budget 750

PHYSICS

International Team Releases Design, Cost for Next Great Particle Smasher

An international team has released a preliminary design and cost estimate for the International Linear Collider (ILC), the hoped-for straight-shot particle smasher that many researchers say is the future of their field.

In Beijing this week, the ILC Global Design Effort (GDE) team reported that the "value" of the 35-kilometer-long machine would be $6.65 billion and 13,000 person-years of labor plus or minus 30%. The value differs from a cost estimate because it does not account for inflation until the machine is completed—in 2016 at the earliest—or so-called contingency to cover potential cost overruns, which different countries handle in different ways, says Barry Barish of the California Institute of Technology in Pasadena, who leads the GDE.

Including such factors would, for example, likely double the amount entered in the ledgers of the U.S. government. So if the United States hosted the machine and bore half the expense, its contribution would total about $7.5 billion, Barish says.

The value estimate is roughly equal to the cost of the Large Hadron Collider (LHC), the 27-kilometer-long circular accelerator under construction at the European particle physics laboratory CERN near Geneva, Switzerland. The LHC cost 4.7 billion Swiss francs (about $3.8 billion), but that does not include the tunnel, which was dug for an earlier machine, or the older accelerators that will feed the LHC.

The LHC and ILC "are in the same ballpark, so I think it is doable," says Albrecht Wagner, chair of the board of directors for the German particle physics laboratory DESY in Hamburg.

But in the end, it's up to the politicians. Some researchers, however, say it is too early to name a price. "I'm afraid that the cost will increase," says Kaoru Yokoya, who leads ILC R&D at the KEK particle physics laboratory in Tsukuba, Japan. "A big increase will kill our project."

The ILC will probe in detail phenomena researchers expect to glimpse at the LHC, which is scheduled to be turned on in November. Many physicists expect the LHC to produce the long-sought Higgs boson and possibly a raft of new particles predicted by a theory called supersymmetry. But the LHC will smash protons into protons, and each proton is itself a tangle of particles called quarks and gluons. So the resulting collisions will be too messy to reveal some key properties of the new particles. The ILC will collide electrons into their antimatter siblings, positrons, to make cleaner collisions that will allow physicists to nail down decay rates and other parameters needed to forge a complete theory.

But first physicists must persuade governments to spring for the machine, and the value estimate is meant to help that process along. In particular, the estimate responds to a request from the U.S. Department of Energy (DOE) to know what the total cost would be. DOE has requested $60 million for ILC R&D in fiscal year 2008.

Over the past year, the GDE transformed a rough "baseline configuration" hacked together from 2 decades of R&D at various labs into a coherent "reference design," making changes that reduced the cost by about 25%. "They have done a good job without stretching the limits beyond reason," says Guenther Grossheke, an accelerator physicist at CERN. "The design has evolved into a cost-conscious design."

The GDE made several changes to the basic layout of the machine. For example, the baseline configuration called for a 7-kilometer circular tunnel near each end of the collider to house accelerators known as cooling rings that would concentrate the electrons and positrons. The new design calls for one circular tunnel near the center that will house both rings. The baseline also split the beams in two to collide particles in the hearts of two detectors; in the new design, the beams collide at just one interaction point, and researchers will swap the gigantic detectors into and out of the line of fire. "Now, there's no way to make it substantially less expensive without reducing the scientific scope of the project," Barish says.

The reference design and value estimate will undergo an international review this spring. That review "is one key to keeping the project from repeating the story of the Superconducting Super Collider, the ill-fated accelerator in Waxahachie, Texas, that Congress killed, uncompleted, in 1993 after its budget exploded from $4 billion to $8 billion. A DOE representative will serve on the review team," says Robin Staffin, associate director for high-energy physics at DOE.

ILC researchers will now use the reference design as the starting point for a more detailed "engineering design." Meanwhile, physicists will lobby for an international framework between governments to fund the project. A decision about where to build the machine is still far off. "My guess is that the international framework will be in place in 2010 at the earliest," says Mitsuaki Nozaki, a particle physicist at KEK and Asian regional director of the GDE. "Until then, we cannot say anything about the final site."

–ADRIAN CHO

Confident. Barry Barish says the price tag for the ILC and its 16,000 accelerating cavities will hold steady.

CREDITS: FERMILAB

ENERGY RESEARCH

BP Bets Big on UC Berkeley for Novel Biofuels Center

Can an oil company take the lead in the biofuels revolution through an unprecedented investment in academic research? BP last week made a 10-year $500 million bet on such a corporate-funded, big-science approach. It chose the University of California (UC), Berkeley, to host its new Energy Biosciences Institute (EBI), a venture that BP's Jim Breson calls "a privately funded national lab." Academic plant scientists and engineers will work alongside BP specialists in the institute, conducting a mix of basic and applied work (see p. 790). A happy partnership, say experts, could yield a new model for industrial support of academic science in service of societal needs. But, as in any marriage, there could be some tensions along the way.

When BP decided to boost its biofuels research agenda, its choices were limited. Having only four biologists on the company's roster of 102,000 employees worldwide ruled out an in-house research effort, says Breson. And Steven Koonin, who stepped down as provost of the California Institute of Technology to become chief scientist at BP in 2004, nixed the idea of simply awarding grants to professors, as a consortium of energy companies led by ExxonMobil is doing (see table). A grants program doesn't foster the "foot traffic, hall traffic, water-cooler talk" that Koonin believes is essential for a successful research enterprise.

The answer, Koonin decided, was private-public, "big science." BP launched a competition to host the institute. UC Berkeley and its partners, neighboring Lawrence Berkeley National Laboratory (LBNL) and the University of Illinois, Urbana-Champaign, edged out four other university teams, says Koonin, in part because of strong connections to biotech, well-equipped scientific facilities, and a history of doing big science. The institute will sit in a yet-to-be-designed building on the UC Berkeley campus, with 25 themed labs. The institute and a smaller sister facility at Illinois will include space for proprietary BP research as well as open labs for academics. A joint board will select projects in what LBNL biochemist Christopher Somerville calls a "spirit of flexibility."

Many operational aspects of the new EBI have yet to be worked out, and some could be contentious. UC Berkeley is no stranger to corporate-funded science as well as the controversy that it can spark. It was criticized for a $25 million, 5-year deal with European pharmaceutical giant Novartis in 1998 that gave the company exclusive licensing provisions for any drug discoveries, and the arrangement was not renewed. Although BP wants its closed areas to be busy with commercial development, Breson emphasizes the company's commitment to sharing intellectual property (IP) with academic scientists and honoring their need to publish. "This has got to be about open research," he says.

But, of course, the devil is in the details. UC Berkeley official Beth Burnside told *Science* that no more than 30% of the total funding would be spent on the 50 BP scientists who will work in proprietary labs at the institute facilities on the partner campuses. Koonin later said that hadn't been finalized and could change yearly. "We're inventing this as we go along," he admits.

A common problem that EBI needs to avoid, says sociologist Lawrence Busch of Michigan State University in East Lansing, who has studied the Novartis deal, is "mismatched expectations" about research payoffs. Breakthrough drugs and lucrative licensing deals "just didn't happen," he warns. One silver lining, Busch says, was that "fears of academic freedom lost or that Novartis was buying the department really didn't pan out." Burnside says UC Berkeley scientists leveraged Novartis funds to do much more molecular biology.

Other scientists stress the importance of building trust through daily scientific interactions to prepare for the inevitable fights over IP and other thorny issues. Microbiologist Jeffrey Gordon of Washington University in St. Louis, Missouri, a veteran partner with Pfizer and other drug companies that Pfizer has acquired, says that the most successful projects flow from joint research proposals reviewed by all the partners. A 2001 decision by computer chip giant Intel to set up a lab adjacent to the UC Berkeley campus raised the productivity of university software researchers, says computer scientist Joseph Hellerstein, because there were suddenly "more smart people to work with." In contrast, he says, a locked room meant for proprietary research "was never used."

Although the type of industry-university collaboration envisioned at EBI "isn't that new," says economist Donald Siegel of UC Riverside, who has studied such arrangements, the scale—a decade-long commitment and a $50 million annual budget—is unprecedented. The magnitude also raises questions about the proper role of industry on campus, says economist David Audretsch of Indiana University, Bloomington. "We need companies to be working with campuses, but we don't want industry-driven universities. This is right on that line."

–ELI KINTISCH

Notable Corporate/University Partnerships

Company	University	Corporate Investment	Time frame	Announced	Focus
BP	UC Berkeley	$500 M	10 years	2007	Biofuels
ExxonMobil, Toyota, Schlumberger, and General Electric	Stanford	$225 M	10 years	2003	Biofuels, renewable energy
Pfizer	Scripps Research Institute	$100 M	5 years	2006	Biomedical research

AFRICAN SCIENCE

African Leaders Endorse Science Initiatives

PRETORIA, SOUTH AFRICA—Science and technology (S&T), although billed as the main themes of last week's African Union (AU) summit in Addis Ababa, Ethiopia, had to wait as leaders addressed pressing issues such as the Darfur conflict. But on the last day, the heads of state approved initiatives aimed at bolstering research and increasing the continent's clout on intellectual-property issues. With little public debate, the leaders urged member states to revitalize their universities and spend at least 1% of gross domestic product (GDP) on research and development by 2010. In addition, the summit called for more extensive S&T agreements with other developing regions, announced new scholarships to stimulate the study of science, and declared 2007 to be Africa's "scientific innovation year."

Upping the ante. Rwanda's President Paul Kagame pledged to double science and technology spending.

"Maintaining the status quo is not an option for Africa," said Rwandan President Paul Kagame. He said Rwanda planned to double its S&T spending to 3% of GDP over the next 5 years and bolster its research institutions. Most African nations are now well below the 1% benchmark. Agronomist John Mugabe, who directs the S&T office of the New Partnership for African Development, which promotes social and economic progress in Africa, says the key point is that "the presidents themselves committed to strengthening science and technology."

One initiative endorsed by Africa's science ministers and tacitly approved by AU leaders is the development of a 20-year biotechnology strategy to channel resources into regional specialties. For example, East Africa, which already has livestock biotech expertise, would be encouraged to build on this base. Other niches include crop biotech in the west, pharmaceutical biotech in the north, medical biotech in the south, and biodiversity research in Central Africa.

Solid-state physicist Phil Mjwara, director-general of South Africa's S&T department, says "there wasn't much discussion" of specifics; details will need to be ironed out by science ministers. He agrees that biotechnology is important but cautions that "it will take a long time to really get it going."

Mjwara and Mugabe agree that creating a pan-African Intellectual Property Organization (IPO)—a concept endorsed by summit leaders—would be an important step toward establishing common standards and goals. Currently, there are two separate organizations on the continent, each representing 16 nations; 20 other countries, mostly in northern Africa, are not represented at all. The heads of state asked for a report by July on how to incorporate existing IPOs into a pan-African IPO that would take up issues such as traditional knowledge and genetic resources.

Not included on the summit agenda, Mugabe says, was a controversial proposal to create a new African science and innovation fund. He expects it will be taken up by science ministers later this year. Its goal would be to pool resources from African nations and outside donors and distribute competitive research and development grants throughout the continent.

"There is a real potential for research synergies" if these agreements lead to pan-African strategies, said Mjwara. But finding common ground on a diverse continent whose 52 nations have widely varying resources and priorities will continue to be a challenge.

—ROBERT KOENIG

SCIENTIFIC WORKFORCE

Racism Allegations Taint Chinese Effort to Recruit Overseas Talent

A new program that aims to inject some international blood into China's scientific workforce has come under fire for excluding ethnic Chinese. Officials at the Chinese Academy of Sciences (CAS) counter that critics are missing the point.

Last September, CAS launched a program to provide fellowships for up to 50 "international young researchers" under the age of 35 to work in CAS laboratories for 1 year with salaries capped at 100,000 yuan, or $12,900, renewable for a second year. The announcement received little publicity until last week, when an anonymous posting on the popular U.S.-based Internet forum *Jiaoyu Yu Xueshu* (Education and Scholarship) noted that the Chinese version of the fellowship rules limits the awards to *fei huaxi waiji* (foreign nationals not of Chinese descent). That phrase sparked a flurry of angry reactions on Chinese Internet sites abroad blasting the program as "racist" and "discriminatory."

The complaints have merit, some experts say. The distinction between *huaxi* (persons of Chinese descent) and *fei huaxi* is an "interesting" form of racism, says Stevan Harrell, an anthropologist at the University of Washington, Seattle, who has done fieldwork in China for more than 20 years. The premise is that "anyone who is of Chinese descent somehow has a connection with China"—even a second- or third-generation American *huaxi*, for instance—and therefore is not considered a "real" foreigner, Harrell explains.

The deputy director of CAS's education department, Li Hefeng, defends the program. Li points out that CAS already has numerous incentive programs to attract *huaxi*, such as the One-Hundred-Talent Program. "About a third of researchers in U.S. and Japanese labs are international," says Li, "but CAS has almost none." Until recently, "we worried whether *fei huaxi* could adapt to our country's work, living, and especially cultural environment," he says. But conditions are improving, Li says, and in the future there may be no need for separate programs for *huaxi* and *fei huaxi*. The "main goal," he maintains, "is to expand international exchange and collaboration." In this case, international means non-Chinese ancestry.

—HAO XIN

CREDIT: CHIP EAST/REUTERS

MEDICINE

Rett Symptoms Reversed in Mice

Some of the dramatic neurological problems of Rett syndrome can be reversed in an experimental mouse model, researchers have found. Although the work does not have direct therapeutic applications, scientists studying the devastating genetic disorder hail the findings as a sign that treatments are at least possible in principle. "This is very exciting," says Huda Zoghbi, a geneticist at Baylor College of Medicine in Houston, Texas. "I guess researchers and the families and patients hope that, as we uncover [biochemical] pathways that could be safely manipulated, we can recover some function in these girls."

Rett syndrome affects roughly one in 10,000 girls. They develop normally for the first 6 to 18 months of life, but they then begin to lose mobility and their cognitive development stalls. Rett girls can live well into adulthood, but they endure severe mental and physical deficits.

Mutations in a gene called *MECP2* are to blame. How mutations in this single gene on the X chromosome, itself a regulator of other genes, causes neurological problems isn't known (*Science*, 8 December 2006, p. 1536), but the new study, published online this week by *Science* (www.sciencemag.org/cgi/content/abstract/1138389), suggests that the damage to the nervous system may not be permanent.

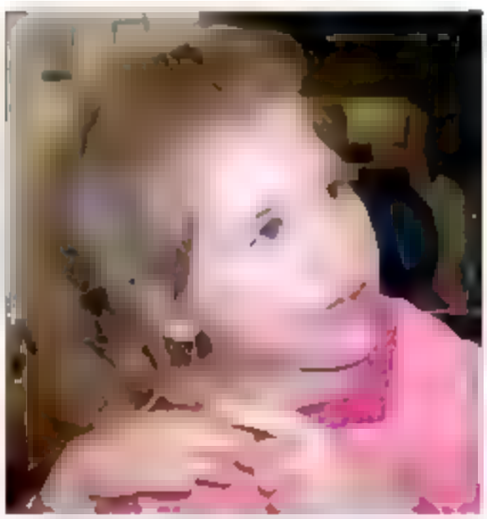
Jacky Guy and Adrian Bird of the University of Edinburgh, U.K., and colleagues created mice with a genetic roadblock—a string of DNA—inserted into *Mecp2* (the mouse version of the human gene) that prevented cells from reading the gene to make the protein it encodes. Female mice with the blocked *Mecp2* developed normally for 4 to 12 months before showing Rett-like symptoms, including impaired mobility, an abnormal gait, tremors, and breathing difficulties.

Then the researchers turned *Mecp2* back on, exploiting another gene they'd bestowed on the mice. This gene encoded a hybrid protein: a DNA-splicing enzyme fused to an estrogen receptor. The enzyme can recognize and remove the roadblock in *Mecp2*, but the attached estrogen receptor prevented it from entering the cell nucleus. By injecting the mice with tamoxifen, a drug that binds estrogen receptors, the researchers sent the hybrid protein scuttling into the nucleus to snip out the roadblock and restore *Mecp2*.

After the mice had received five weekly tamoxifen injections, the Rett-like symptoms all but disappeared. A few of the rodents continued to walk with their hindlimbs abnormally far apart, but otherwise they were hard to distinguish from their genetically normal relatives. It was a pleasant surprise, because researchers had feared that the developmental loss of *Mecp2* led to missing or permanently disabled neural connections. "The general perception is that once the brain has missed out big time on some ingredient of normal development, it's never going to be able to recover," Bird says. "We thought maybe we'd get amelioration of the symptoms, but we didn't anticipate that things would be reversed on the scale that we found."

The findings suggest that the lack of *Mecp2* doesn't do irreversible damage to neurons, says Rudolf Jaenisch of the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts. "It's almost like a dream result," he adds.

Still, the study doesn't point to an obvious strategy for treating Rett syndrome. Human gene mutations can't be repaired by the technique the Edinburgh team used, and simply boosting MECP2, either by gene therapy or administering the protein, is not likely to work, Zoghbi and others say. That's because girls with Rett syndrome already make MECP2 protein in about half their cells, thanks to a good copy of *MECP2* on their second X chromosome. These cells would end up with a surplus of MECP2, which appears to be just as damaging as a deficit is. Finding alternative strategies won't be easy, but the mouse work suggests that such efforts are well worth pursuing. —GREG MILLER



CREDIT: ALAN PERCY/UNIVERSITY OF ALABAMA, BIRMINGHAM

Hope. Girls with Rett syndrome have impaired mobility; mouse studies hint at a remedy.

SCIENCESCOPE

France: The Blame Game

PARIS—A new government report blames badly managed public research for historically poor commercial payoffs and calls for funding projects rather than institutions. "Despite measures taken since the 1999 innovation and research law, the commercialization of research has not progressed in France for 15 years," the report says.

Bertrand Monthubert, president of the advocacy group Sauvons La Recherche, says the 250-page report "takes no account of the size of labs or the fact that most public contracts are awarded for priority research areas, such as nanotechnologies or energy." Monthubert says industry is to blame and that companies should hire more of the 10,000 Ph.D. students who graduate French schools each year. Research unions are citing the report in a call to scientists to join civil servants in street protests this week. —BARBARA CASASSUS

FDA Reports Drug Studies Backlog

A yearly update from the U.S. Food and Drug Administration (FDA) on whether companies are monitoring the safety and efficacy of therapies after they hit the market shows that few such studies are ongoing: 274 in all, or 17% of the number promised. Meanwhile, nearly two-thirds, or 1026, postmarketing studies, haven't yet started. FDA can't say how many of those are stalled or how many are soon to begin.

Some outsiders say clearer data are needed and that FDA and drug companies should more aggressively pursue the studies, such as those assessing the safety of a certain drug dose. "There's not been a huge amount of progress" in clearing the backlog, says epidemiologist Bruce Psaty of the University of Washington, Seattle. FDA says that some of the $11.2 million requested by the White House this week for drug safety initiatives will go toward postmarketing surveillance. —JENNIFER COUZIN

MIT Averts Hunger Strike

The Massachusetts Institute of Technology in Cambridge agreed this week to let an African American stem-cell researcher who did not receive tenure remain on campus several additional months. James Sherley threatened in December to begin a hunger strike over the tenure denial that he attributes to racism. The university, which said this week that racism was not a factor in the tenure decision, plans to set up a committee to examine the status of minority faculty members on campus.

—ANDREW LAWLER

2008 U S BUDGET

Research Rises—and Falls—in The President's Spending Plan

Just as he has stayed the course in Iraq, President George W. Bush has stuck to his guns with his budget proposals. On 5 February, he sent Congress a 2008 budget request for science that favors a handful of agencies supporting the physical sciences and puts the squeeze on most of the rest of the federal research establishment as part of an overall $2.9 trillion plan that clamps down on most civilian spending.

Last year, the president's budget proposed the American Competitiveness Initiative (ACI), a 10-year doubling of the National Science Foundation (NSF), the Department of Energy's (DOE's) Office of Science, and core labs at the National Institute of Standards and Technology (NIST). At the same time, he asked Congress to cut the National Institutes of Health (NIH), which makes up almost half the $60 billion federal science and technology budget. His 2008 request follows suit. If it were to be adopted intact—a wholly unlikely prospect, given the Democratic control of Congress—government support for basic research would rise by a minuscule 0.5% from his 2007 request.

Accordingly, the new request is getting a decidedly mixed reception from science groups. "The [president's] budget has much positive news for the nation's research universities, but it also raises some very serious concerns," says Robert Berdahl, president of the 62-member Association of American Universities (AAU). He praises the ACI funding but chastises the Administration for "shortchanging" defense research and making what amounts to a $500 million cut in NIH funding. Democratic lawmakers who follow science are equally ambivalent. "The president's budget includes a few good targets for R&D funding but ignores too many of our country's priorities," says Representative Bart Gordon (D-TN), chair of the House Committee on Science and Technology.

Biomedical lobbyists don't see much of a silver lining. "It's distressing," said Jon Retzlaff, legislative chief at the Federation of American Societies for Experimental Biology, which has called for a doubling of NIH. "Our goal will be to have this changed when it goes through Congress."

The 2008 budget is harder to interpret than most presidential requests because Congress hasn't finished work on the 2007 budget. The previous, Republican-led Congress adjourned in December after ordering agencies to keep spending at 2006 levels until 15 February, and the new Democratic majority is now putting together a final 2007 budget. A resolution adopted on 31 January by the House would hold most federal agencies to 2006 levels, but with a few notable exceptions, including NSF, DOE's Office of Science, and NIST. The Senate is expected to take up the measure shortly.

The 2008 request for science will play out in the midst of the bitter debate over the Iraq war and efforts to shrink the federal deficit. Here are some individual agency highlights.

NIH: Its 2008 request pays special attention to younger scientists despite containing less than the amount Congress is expected to approve for 2007. New and competing grants in the president's plan would surpass 10,000 for the first time in years, says NIH Director Elias Zerhouni, on the way toward achieving "stable, sustainable" growth. Funding for high-priority "Roadmap" projects under the NIH director's control will increase, as will special grants for first-timers.

Biomedical lobbyists figure that NIH would actually receive $500 million less under the president's budget than in 2007, in part because its contribution to the Global Fund to Fight AIDS, Tuberculosis, and Malaria and other diseases in the developing world would rise from $99 million to $300 million. They also note that the total number of grants would drop. Asked about the likelihood of such a cutback, a Senate majority staffer said, "We're not going to let that happen."

NSF: Director Arden Bement says he's "more than happy" for a budget that has room for a $82 million initiative to add computational elements into the biological and physical sciences and engineering as well as a $24 million increase in a $90 million program for major research infrastructure at universities with limited endowments. NSF hopes to add 200 graduate research fellowships in 2008.

With regard to major new facilities, the agency has requested $32 million to

Continued on p. 753

A Science Budget Up for Grabs (in $ Millions)

Agency	House 2007 CR*	2008 request	% change
National Institutes of Health	28,931	28,621	−1.1%
Transfer to Global Fund	99	300	+203%
National Science Foundation	5,916	6,429	+8.7%
Research†	4,766	5,131	+7.7%
Education	700	750	+7.1%
NASA	16,247	17,309	+6.5%
Science	5,251	5,516	+5.0%
Exploration	3,401	3,924	+15.4%
Department of Energy, Office of Science	3,796	4,397	+15.6%
Department of Defense basic research	1,564	1,428	−8.7%
DARPA basic research	145	153	+5.5%
Homeland Security, science and technology	1,003	976	−2.7%
Department of Commerce			
NOAA oceanic and atmospheric research	370	358	−3.2%
NIST labs	426	492	+15.5%
Advanced Technology Program	79	0	
Environmental Protection Agency science	563	540	−4.1%
USDA competitive research	190	257	+35.3%
U.S. Geological Survey	978	975	0.3%
Food and Drug Administration‡	1,965	2,085	+6.1%

* Passed by the House 31 January, awaits Senate action.
† Includes transfer of EPSCoR program from education to research account.
‡ Includes user fees.

Pick a number. The current House spending plan may be the best yardstick to measure the president's request for 2008, because Congress hasn't completed work on the 2007 budget.

SOURCE (BOTTOM): OFFICE OF MANAGEMENT AND BUDGET; INDIVIDUAL AGENCIES

Highlights From the Budget

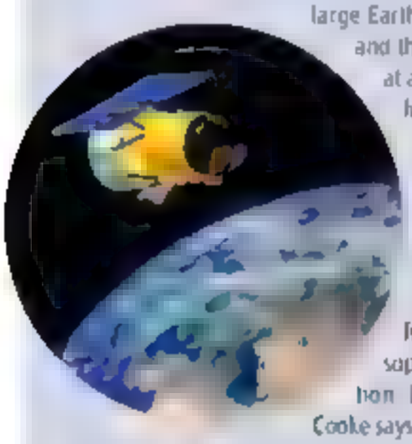
NASA Tries to Make Space for Science

NASA came under heavy fire last month from the National Academies' National Research Council for cutting important earth science missions, and just this week, another academy panel complained that overruns in large projects such as the James Webb Space Telescope, slated for a 2013 launch, are hurting smaller astrophysics efforts. But NASA officials say that their 2008 budget will put the agency's beleaguered $5.5 billion science program back on a more balanced track.

Exhibit A, say agency officials, is the proposed restoration of several projects hanging in limbo during the past year. A joint U.S.-Japanese spacecraft designed to measure global precipitation will be launched in two phases in 2013 and 2014. And after several years of overruns and delays, two smaller astrophysics missions—Kepler, designed to discover extrasolar planets, and the Wide Field Infrared Survey Explorer (below), which will produce a detailed map of the infrared universe—will by next year be given the money they need to fly.

That news may not be enough to satisfy astrophysicists worried that the queue of missions is quickly emptying. "There are no low cost, quick-response science programs being prepared today," says Martha Haynes, an astrophysicist at Cornell University and vice chair of the academy panel that released its report on 7 February. NASA requested the report last year.

It is clear, however, that NASA's science program remains in crisis after having to pony up $2.44 billion from its 2007–2011 budget plan to cover shuttle and space station shortfalls. No spacecraft are slated to follow the large Earth-observing platforms now in orbit, and the earth sciences budget will remain at about $1.5 billion for the foreseeable future. Several important astrophysics flights, such as the Space Interferometry Mission, remain on hold because of budget constraints.



Even lunar science, now in favor because of its connection to human exploration, faces an uncertain future after the projected cost of a few sophisticated rovers doubled to $1.5 billion. Deputy exploration chief Douglas Cooke says such rovers may not be necessary at all. "We need to have a really good map," he says. "It's hard to say we need much more than that." Nevertheless, NASA Ames Research Center in Mountain View, California, will put together a cheaper and faster option for the rovers.

—ANDREW LAWLER

NIH Children's Study Sparks Fight

Director Elias Zerhouni says the National Institutes of Health (NIH) can't afford it. But legislators believe that a massive search for the sources of chronic ailments such as asthma and diabetes is worth $3 billion over the next 25 years.

This week, for the second straight year, President George W. Bush's budget request to Congress zeroed out funding for the National Children's Study, which would enroll 100,000 newborns (beginning in the womb) and monitor their health and exposure to environmental risks for a quarter century (*Science*, 10 December 2004, p. 1883). Authorized by Congress in 2000, the study received $12 million in 2006 to support seven "vanguard" centers that will help design the study. Last week, a House spending plan for 2007 included $69 million for the study, and supporters are gearing up to push for $111 million in the 2008 fiscal year that begins on 1 October.

Why does this study have such strong legislative backing despite NIH's plea of poverty? Supporters argue that it would more than pay for itself if it leads to even a 1-year 1% reduction in the key diseases it examines. But one biomedical lobbyist who requested anonymity claims that advocates created a large and geographically diverse constituency by naming the vanguard centers and identifying nearly 100 future study sites. The lobbyist, who called the tactics an "outrage," likened them to methods used by any special interest group to retain federal funding for a large project.

The children's study is not a special interest project, says lobbyist John Porter, a partner at the Hogan & Hartson law firm in Washington, D.C. Porter, a former Republican representative from Illinois who chaired the House panel that sets NIH's budget, has been hired by the vanguard centers. "We respected the fact that funding is very tight and difficult for investigators," he says, adding that Congress has put the study in the NIH director's budget rather than having it compete against investigator-initiated grants at individual institutes.

The study's cost "is not trivial," concedes Leonardo Trasande, a preventive medicine researcher who has championed it alongside his colleague Philip Landrigan at Mount Sinai School of Medicine in New York City. But he says it's the best way to find the complex factors that are leading an "epidemic" of chronic childhood illness.

—ELIOT MARSHALL

NSF Education Program Rebounds

The Bush Administration's efforts to rein in the National Science Foundation's (NSF's) education programs may be coming to an end, with an assist from a congressionally mandated review of existing federal efforts to improve math and science education.

NSF's 2008 budget includes $30 million to revive its Mathematics and Science Partnership program, a competitive grants program that teams universities and local school districts to improve elementary and secondary school math and science. Begun in 2002 as a $200 million presidential initiative, the NSF program had shrunk to a proposed maintenance-level $46 million in 2007 while a similar program in the Department of Education that awards block grants to every state grew rapidly (*Science*, 24 February 2006, p. 1092). The Administration's request would permit a new round of competitive 5-year awards, says NSF Director Arden Bement, along with developing "tool kits" for teachers based on successful programs to date.

Many legislators have long argued that NSF could do a better job than the Department of Education, and a review of all federal programs by the Interagency Academic Competitiveness Council (ACC) to be released later this month seems to bear them out. "I think that the ACC recognizes the quality of our assessments and our success in achieving significant student outcomes," says Bement, who has seen a draft of the upcoming report. Robert Shea, associate director for management at the White House Office of Management and Budget, told *Science* last month that NSF is among those agencies "with a strong track record for rigorous evaluation" that could be models for other agencies trying to improve science and math instruction.

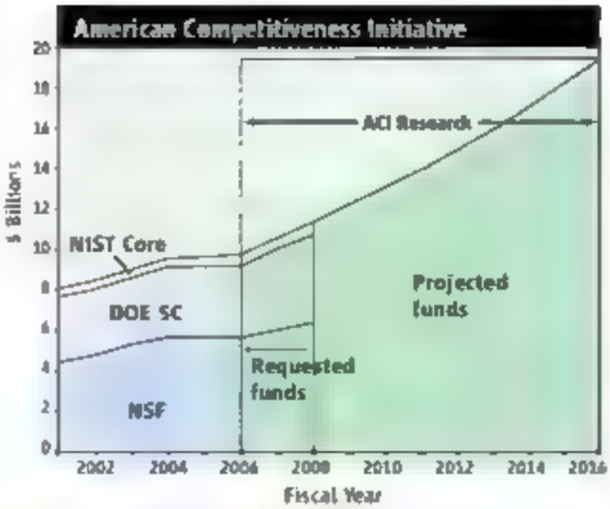
—JEFFREY MERVIS

CREDITS: TOP TO BOTTOM: SCIENCE FACTORY/GETTY IMAGES; JPL/NASA

begin upgrading the Laser Interferometer Gravitational Wave Observatory, and it has stretched out the scheduled ramp-up of national environmental and ocean observatory networks because of continued tweaking of their designs. Bement says he also hopes to win congressional approval in 2007 to begin building a $120 million Arctic research vessel.

HOMELAND SECURITY: The Department of Homeland Security (DHS), whose Science and Technology directorate was reprimanded by Congress last year for mismanaging its finances, wants to cut its university programs by 20%, to $39 million. And although officials also plan to add four university-based centers of excellence to its current roster of seven, that expansion will likely translate into less money for existing centers. Jay Cohen, the directorate's new undersecretary, promises that the program will bounce back. "Please view this as a transition year as we realign the [centers] program with our mission," he says, adding that DHS remains "fully committed to basic research."

DEFENSE. The silver lining in an 8.7% cut for basic research at the Department of Defense (DOD) is a big boost to the National Defense Education Program, from $19 million to $44 million. The program, begun in 2006, provides scholarships to U.S. undergraduate and graduate students in science and engineering disciplines related to national defense in hopes of replenishing DOD's aging science and technology employees. "The idea is definitely resonating with folks on the Hill," says AAU's Matthew Owens. The basic research portfolio at the Defense Advanced Research Projects Agency would rise by 5.5%, to $153 million.



Still competitive. The Administration has requested a second year of big increases for the three agencies under ACI.

ENERGY: Undersecretary of Science Raymond Orbach says that a second consecutive year of large increases will allow DOE to take "our first step into new territory." The expected 2007 increase, he adds, "repaired the damage" to programs long starved for funds.

The 2008 request would give $160 million toward the $6 billion International Thermonuclear Experimental Reactor in Cadarache, France. Projects related to applied energy science would enjoy a whopping 26% increase overall above the 2007 request, notably for biomass research ($213 million) and nuclear waste recycling studies ($95 million).

SPACE. NASA Administrator Michael Griffin says he is happy with the proposed 3.1% increase over the president's 2007 request. But the 2007 budget coming down the pike from the new Democratic-run Congress is quite another matter. "We're on the receiving end of a budget we don't like," grouses Griffin, referring to a half-billion-dollar reduction in NASA's effort to build a new launcher to take humans to the moon. That cut, if it holds up, could greatly extend completion of the system, now planned for 4 years after the shuttle's last flight in 2010.

NIST: The proposed 21% boost over the House-approved 2007 funds would advance a proposed lab expansion at the agency's Boulder, Colorado, campus and an upgrade to the NIST Center for Neutron Research in Gaithersburg, Maryland. It would also provide $22 million for five new programs in areas such as nanotechnology and climate science. "This is actually a great day for NIST," says the agency's director William Jeffrey. As in previous requests, the president would zero out the Advanced Technology Program (ATP), which helps companies commercialize nascent technologies. The House spending bill offers ATP a 1-year reprieve, at $79 million.

EPA: The Environmental Protection Agency's Office of Research and Development would suffer a 4.1% cut from the House-approved level. That's less than the 6.7% reduction the Administration sought for 2007. EPA chief scientist George Gray says that the 2008 request includes small boosts for intramural research on nanotechnology and risk assessment.

NOAA. The National Oceanic and Atmospheric Administration is seeking $40 million more for ocean research as part of an overall $123 million boost under the president's Ocean Action Plan (*Science*, 2 February, p. 585). But NOAA's main program for research, the Office of Oceanic and Atmospheric Research (OAR), would drop by 3.3% from the House 2007 spending bill, which hews to 2006 levels.

AGRICULTURE: The Administration reprised its goal of boosting competitive research at the U.S. Department of Agriculture. The National Research Initiative would see a 38% jump to $257 million, just slightly less than the agency requested last year. And the White House restated its proposal to create competitive awards for 9.4% of the so-called formula funding awarded to land-grant universities each year. Agriculture school lobbyists say they will continue to oppose that move unless the overall pie gets bigger. That doesn't happen in the Administration's request for $530 million, at roughly the 2006 level.

–JEFFREY MERVIS

Reporting by Yudhijit Bhattacharjee, Adrian Cho, Constance Holden, Eli Kintisch, Andrew Lawler, Eliot Marshall, Robert F. Service, and Erik Stokstad.

Science Adviser Battles Cancer

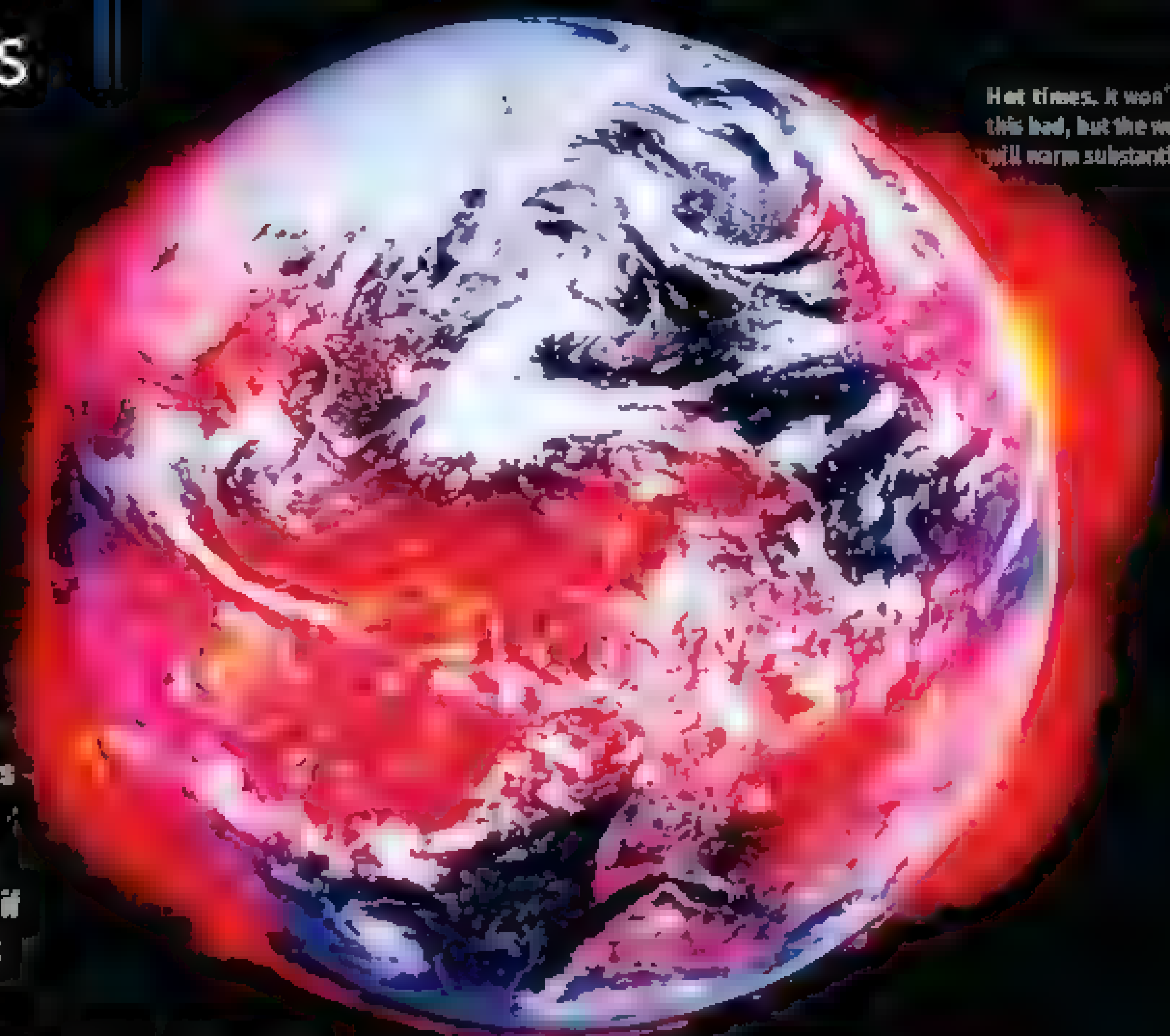
Last month, John Marburger became the longest serving presidential science adviser in U.S. history. The 66-year-old laser physicist has been at the helm for many policy battles in the 5 1/2 years he's served President George W. Bush, but he's currently engaged in an even tougher fight—against cancer.

In November, Marburger was diagnosed with non-Hodgkin's lymphoma and began a round of treatments. The effects of the chemotherapy have forced the former president of Stony Brook University and one-time director of Brookhaven National Laboratory to work much of the time from his home in New York.

This week, Marburger's condition became public when he missed the Administration's annual science budget briefing, over which he normally presides. "Jack can't be here because he is currently fighting a battle with cancer, which he is winning," explained Richard Russell, deputy director for technology and Marburger's longtime aide. Marburger is expected to testify next week before the House science committee.

–J.D.M.

SOURCE: OSTP



Hot times. It won't get this bad, but the world will warm substantially.

They've said it before, but this time climate scientists are saying it with feeling: The world is warming; it's not all natural, it's us; and if nothing is done, it will get a whole lot worse

Scientists Tell Policymakers We're All Warming the World

THE LAST TIME THE INTERGOVERNMENTAL Panel on Climate Change (IPCC) assessed the state of the climate, in early 2001, it got a polite enough hearing. The world was warming, it said, and human activity was "likely" to be driving most of the warming. Back then, the committee specified a better-than-66% chance—not exactly a ringing endorsement. And how bad might things get? That depended on a 20-year-old guess about how sensitive the climate system might be to rising greenhouse gases. Given the uncertainties, the IPCC report's reception was on the tepid side.

Six years of research later, the heightened confidence is obvious. The warming is "unequivocal." Humans are "very likely" (higher than 90% likelihood) behind the warming. And the climate system is "very unlikely" to be so insensitive as to render future warming inconsequential.

This is the way it was supposed to work, according to glaciologist Richard Alley of Pennsylvania State University in State College, a lead author on this IPCC report. "The governments of the world said to scientists, 'Here's a few billion dollars—get this right,'" Alley says. "They took the money, and 17 years after the first IPCC report, they got it right. It's still science, not revealed truth, but the science has gotten better and better and better. We're putting CO_2 in the air, and that's changing the climate."

With such self-assurance, this IPCC report may really go somewhere, especially in the newly receptive United States (see sidebar, p. 756), where a small band of scientists has long contested IPCC reports. Coordinating lead author Gabriele Hegerl of Duke University in Durham, North Carolina, certainly hopes their report hits home this time. "I want societies to understand that this is a real problem, and it affects the life of my kids."

Down to work

Created by the World Meteorological Organization and the United Nations Environment Programme, the IPCC had the process down for its fourth assessment report. Forty governments nominated the 150 lead authors and 450 contributing authors of *Climate Change 2007: The Physical Science Basis*. There was no cabal of senior insiders: 75% of nominated lead authors were new to that role, and one-third of authors got their final degree in the past 10 years. Authors had their draft chapters reviewed by all comers. More than 600 volunteered, submitting 30,000 comments. Authors responded to every comment, and reviewers certified each response. With their final draft of the science in hand, authors gathered in Paris, France, with 300 representatives of 113 nations for 4 days to hash out the wording of a scientist-written Summary for Policymakers.

The fact of warming was perhaps the most straightforward item of business. For starters, the air is 0.74°C warmer than in 1906, up from a century's warming of 0.6°C in the last report. "Eleven of the last twelve years rank among the 12 warmest years in the [156-year-long] instrumental record," notes the summary (ipcc-wg1.ucar.edu). Warming ocean waters, shrinking mountain glaciers, and retreating snow cover strengthened the evidence.

CREDIT: ROB ATKINS/THE IMAGE BANK/GETTY IMAGES

So the IPCC authors weren't impressed by the contrarian argument that the warming is just an "urban heat island effect" driven by increasing amounts of heat-absorbing concrete and asphalt. That effect is real, the report says, but it has "a negligible influence" on the global number. Likewise, new analyses have largely settled the hullabaloo over why thermometers at Earth's surface measured more warming than remote-sensing satellites had detected higher in the atmosphere (*Science*, 12 May 2006, p. 825). Studies by several groups have increased the satellite-determined warming, largely reconciling the difference.

This confidently observed warming of the globe can't be anything but mostly human-induced, the IPCC finds. True, modeling studies have shown that natural forces in the climate system—such as calmer volcanoes and the sun's brightening—have in fact led to warming in the past, as skeptics point out. And the natural ups and downs of climate have at times warmed the globe. But all of these natural variations in combination have not warmed the world enough, fast enough, and for long enough in the right geographic patterns to produce the observed warming, the report finds. In model studies, nothing warms the world as observed except the addition of greenhouse gases in the actual amounts emitted.

From studies of long-past climate, including the famous hockey-stick curve of the past millennium's temperature (*Science*, 4 August 2006, p. 603), the IPCC concludes that the recent warming is quite out of the ordinary. "Northern Hemisphere temperatures during the second half of the 20th century were *very likely* higher than during any other 50-year period in the last 500 years," the report concludes, "and *likely* the highest in at least the past 1300 years."

Contrarians have conceded that greenhouse gases may be warming the planet, but not by much, they say. The climate system is not sensitive enough to greenhouse gases to overheat the globe, they say. For the first time, the IPCC report directly counters that argument. Several different lines of evidence point to a moderately strong climate sensitivity (*Science*, 21 April 2006, p. 351). The eruption of Mount Pinatubo in 1991 thickened the stratosphere haze layer and cooled climate, providing a gauge of short-term climate sensitivity. Paleoclimatologists have determined how hard the climate system was driven during long-past events such as the last ice age and how much climate changed then. And models have converged on a narrower range of climate sensitivity.

The IPCC concludes that both models and past climate changes point to a fairly sensitive climate system. The warming for a doubling of CO_2 "is *very unlikely* to be less than 1.5°C," says the report, not the less than 0.5°C favored by some contrarians. A best estimate is about 3°C, with a likely range of 2°C to 4.5°C.

What next?

Looking ahead, the report projects a warming of about 0.4°C for the next 2 decades. That is about as rapid as the warming of the past 15 years, but 50% faster than the warming of the past 50 years. By the end of this century, global temperatures might rise anywhere between a substantial 1.7°C and a whopping 4.0°C, depending on the amount of greenhouse gases emitted. In some models' projections, late-summer Arctic sea ice all but disappears late in this century. It is very likely that extremes of heat, heat waves, and heavy precipitation events will continue to become more frequent. Rain in lower latitudes will decrease, leading to more drought.

On some hot topics, the IPCC comes down on the conservative side. It sees evidence of more intense hurricane activity in the North Atlantic, something many researchers contest, but paints a murky picture elsewhere, in line with doubters' reservations (*Science*, 10 November 2006, p. 910). As to the so-called meridional overturning circulation (MOC), the conveyor belt of currents that delivers warm water to the far North Atlantic, there is not enough evidence to say whether it has slowed under global warming, according to the IPCC, contrary to a high-profile report of a 30% slowing (*Science*, 17 November 2006, p. 1064). But

Warmer, then hot. A middle-of-the-road scenario calls for warmings of more than 6°C in high northern latitudes.

Long-range forecast. IPCC scientists in Paris warn of greenhouse warming getting out of hand.

CREDITS: (TOP TO BOTTOM) IPCC; CHARLES PLATIAU/REUTERS/LANDOV

U.S. POLICY: A PERMANENT SEA CHANGE?

This time around, a stormy yet another international assessment of the planet's climatic health (see main text) might look like throwing gasoline on a fire. Even in the United States, where skepticism about human-induced climate change has long dominated government policy, public concern about global warming was already as high as it ever had been. U.S. media had been at fever pitch on climate for a year and more. And local, state, and national politicians from both of the country's major political parties were pushing or even implementing their own proposals for reining in greenhouse gas emissions.

"It takes a sudden jolt sometimes before we become aware of a danger," former vice president Al Gore says in his Oscar-nominated global warming documentary *An Inconvenient Truth*. If so, the shock to public and political perceptions this time may have come in large part from Mother Nature. There has been "a chronic drip of [media] stories about weather effects that are hard to control," notes geoscientist Michael Oppenheimer of Princeton University, from raging hurricanes such as Katrina and melting Arctic sea ice to mid-January daffodil blooms in Washington, D.C.

Will the worst weather drive the body politic to eventual action on global warming? Some observers believe so. "Probably it's robust," says Oppenheimer. The sentiment toward U.S. action "is just not going to go back." But many of the jolting climate events may themselves go away, at least temporarily. Ice-melting warmth in the Arctic and surges of deadly hurricanes in the Atlantic, among other climate trends, are also subject to natural swings that could temporarily slow or even pause some of global warming's more dramatic effects.

The first American surge in attention to global warming started with a similarly combustible mix of climate science and weird weather. On one of the hottest days of 1988, a muggy Washington, D.C., with drought gripping much of the American West and huge wildfires raging through a tinderdry Yellowstone National Park, leading climate researcher James Hansen of NASA's Goddard Institute for Space Studies in New York City testified on Capitol Hill. This was greenhouse warming, he told Congress confidently.

News coverage, at least, took off, according to a new account by political scientists Maxwell Boykoff of the University of Oxford and Jules Boykoff of Pacific University in Forest Grove, Oregon. It peaked again in 1992, when the first President George Bush signed the United Nations Framework Convention on Climate Change. Then media attention promptly plummeted. It perked up only briefly, the Boykoff brothers note, in response to political events, such as when President George W. Bush rejected the Kyoto Protocol limiting greenhouse gas emissions in 2002.

Nevertheless, global warming never entirely vanished from the American consciousness. In the past couple of years, the whole issue has moved up the agenda, again, says political communications researcher Matthew Nisbet of American University in Washington, D.C., and it did it without the typical boost from a single, triggering political event. The public still hasn't elevated global warming to its highest levels of environmental concern, Nisbet notes, but the subject has permeated television, movies, and books, segued onto the business, style, and gardening pages, and broken into daily conversation. Media attention reached an all-time high in 2006, Nisbet has found, as gauged by the number of articles in elite newspapers. And the mutualistic relation between the media and politicians was cranked way up as reporters fed off the policy debate and politicians drew strength from media coverage.

So what got the pot boiling so high in the United States this time around? Many observers see global warming moving to the front burner much the way

the IPCC goes on to project a very likely reduction in MOC flow by the end of the century, perhaps on the order of 25%. Contrary to the climate catastrophe movie *The Day After Tomorrow*, however, a slowing of the MOC would not freeze up the North Atlantic. The region won't even cool off, thanks to greenhouse warming. And it's very unlikely the MOC will abruptly shut down this century, the report says.

The IPCC is overly conservative, in the opinion of some newly outspoken scientists, when it comes to the fate of the world's great ice sheets—on Greenland and Antarctica—and the likely rise in sea level. The facts are not in much dispute. The ocean is warming and therefore expanding; mountain glaciers are melting into the sea; and Greenland is melting around its edges as well. That drove up sea level as fast as 3 millimeters per year lately. The IPCC projects that sea level will continue to rise 28 to 43 centimeters in this century, depending on emissions.

It is also generally agreed that the IPCC calculation leaves out a potentially important factor. Some glaciers draining ice from Greenland and West Antarctica have sped up in the past 5 to 10 years, some of them doubling their speed (*Science*, 24 March 2006, p. 1698). But this glacier acceleration is not included in the IPCC sea-level projection "because a basis in published literature is lacking," according to the report.

That didn't sit well with some researchers.

CREDIT: NICOLE FRUGE/AP

ozone destruction did in the 1980s. Theoretical predictions were prompting some governments to begin to curtail chlorofluorocarbon (CFC) emissions. Then researchers recognized the springtime ozone hole hovering over the Antarctic, galvanizing international negotiations on eliminating CFCs.

Global warming may never have the equivalent of the ozone hole, but the cumulative effect could be the same. Americans "are starting to see changes in the weather," says Eileen Claussen, president of the Pew Center on Global Climate Change in Arlington, Virginia. There was the 2004 hurricane season, with four hurricanes wreaking havoc across Florida, followed by Katrina in 2005, year after year of record-breaking shrinkage of Arctic sea ice, accompanied by images of hungry polar bears, glaciers accelerating their rush to the sea in Greenland and Antarctica, driving up sea level, and those daffodils in the nation's capital.

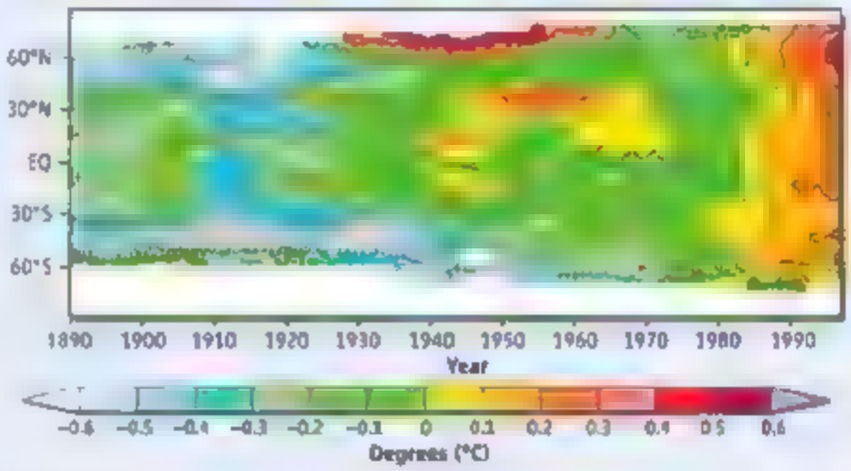
Political and economic factors have also helped fuel the fire. Many alternatives to expensive oil, for example, would also ease greenhouse warming. But to the extent that dramatic climate events have heightened interest, global warming activists have a sometimes unreliable helper, researchers note. The climate system swings from warm to cool and back, from wet to dry. El Niño's unusual warmth in the tropical Pacific is just one of many natural climate variations. And the climate system does it all quite on its own.

Such natural variability could temporarily reinforce or reverse global warming effects. In a study in press at the *Journal of Climate*, for example, modeler Gabriele Hegerl of Duke University in Durham, North Carolina, and colleagues reconstructed past Northern Hemisphere temperature from records such as tree rings and then apportioned the warming using a simple climate model. They found that rising levels of greenhouse gases account for about one-third of the large, rapid warming over the first half of the past century, but another third must have been a natural warming.

What warms the hemisphere can also cool it, according to a modeling study reported 30 January in *Geophysical Research Letters*. Modeler Rong Zhang and his colleagues at the Geophysical Fluid Dynamics Laboratory in Princeton, New Jersey, found that oscillations in the flow of warm currents into the North Atlantic could have added substantially to the early-20th century warming, contributed to the mid-century pause in warming, and bolstered the obvious greenhouse-fueled warming of the past few decades. If what has been going up and down for at least a century goes down again, the ongoing warming that took off in the '70s could noticeably slow in the next 10 or 15 years.

Natural variability is even stronger in places such as the Arctic (*Science*, 5 January, p. 36), one of two climatic "canaries in the coal mine" Gore cites in his film. In the 1930s to 1940s, the Arctic warmed to even higher temperatures than now, only to cool back down by the 1970s. Drawing on 12 climate models, Arctic researcher James Overland of the National Oceanic and Atmospheric Administration's Pacific Marine Environmental Laboratory in Seattle, Washington, sees the recent rapid Arctic warming "as a fairly strong natural variability signal on top of long-term [human-made] change. It's very likely we could have a 5-year period of colder temperatures, and people could say, 'Aha, we don't have global warming.' "

Other worrying climate trends could pause or moderate as well. Part of the surge in powerful Atlantic hurricanes since 1995 is attributable to a natural cycle in the proportion of major storms, according to meteorologist Gregory Holland of the National Center for Atmospheric Research in Boulder, Colorado (*Science*, 10 November 2006, p. 910). So the Atlantic could quiet down a bit, he says, although perhaps not until 2020 or later. Another climate threat—the collapse of the climate-moderating currents of the Atlantic—seems to have receded already. Late in 2005, oceanographers reported measuring a sizable 30% slowdown in the so-called meridional overturning circulation only to concede late last year that their record was so noisy that they couldn't reliably detect any change after all (*Science*, 17 November 2006, p. 1064).



Warm and warm again. The world, especially the Arctic, warmed in the 1930s and '40s (warm colors), in large part due to natural variability.

Just how natural climate variability will interact with political divisiveness, the public's mood swings, and the cyclic economics of energy is unclear at this point. Some say, however, that the mix of bizarre weather and politics boosted by the media now ensures there's no turning back. Even if the climate craziness fades for a while, "I think national legislation is inevitable in 4 years," says Claussen. Others, however, think such confidence may be misplaced. Science historian Naomi Oreskes of the University of California, San Diego, recalls the energy craze of the late 1970s, when soaring oil prices drove dreams of energy independence through conservation and alternative fuels. That passed as soon as prices fell. "We got excited for a while, but we didn't take the serious steps."

–R.A.K.

CREDIT: RICHARD ALLEY/PENNSYLVANIA STATE UNIVERSITY

such as Stefan Rahmstorf of the Potsdam Institute for Climate Impact Research in Germany. He authored a *Science* paper last month that extrapolated from the recent sea-level rise to a rise ranging from 0.5 meter to a near-disastrous 1.4 meters by the end of the century. Days before the 2 February IPCC report release, he and others—call them counter-contrarians—spoke out in news reports. The IPCC sections on sea-level rise are "obviously not the full story because ice-sheet decay is something we cannot model right now but we know it's happening," Rahmstorf told the Associated Press. "A document like that tends to underestimate the risk." And the day before the IPCC release, a second paper—co-authored with seven colleagues—was published online by *Science* (www.sciencemag.org/cgi/content/abstract/1136843). The seas have been rising at the uppermost rate projected in past IPCC reports, the authors noted. Sea level "may be responding more quickly than climate models indicate," they wrote.

Such publicly expressed concerns are likely to become more common, says climate modeler Michael MacCracken of the Climate Institute in Washington, D.C. Now that the contrarians have been dealt with, he says, scientists no longer concerned about appearing to be alarmist will be speaking out about the IPCC being overly cautious.

–RICHARD A. KERR

With additional reporting by Michael Balter in Paris.

Wired for sound. Emotions will be monitored at McGill's [illegible]

PROFILE BRAIN MUSIC AND SOUND RESEARCH CENTER

Study of Music and the Mind Hits a High Note in Montreal

After battling to have their field taken seriously, two musician-scientists have founded an interdisciplinary center to understand how—and perhaps why—humans make music

MONTREAL, CANADA—It's Saturday night in Montreal's Latin Quarter and the flamenco dancers are in full whirl at the Casa Cubana restaurant. A tall woman with a gardenia in her hair stomps the wooden floor in staccato bursts as loud as firecrackers, while her partner yelps out a lull-throated *cante*. At a table nearby, Isabelle Peretz sips a glass of Rioja with a wistful smile. "This music always gives me an intense frisson," she says.

But music is more than just a Saturday night pleasure for Peretz, a neuropsychologist at the University of Montreal (UM). Peretz has devoted her career to understanding why and how the human brain allows us to create and respond to sequences of sounds. In 2005, together with brain imaging specialist Robert Zatorre of McGill University in Montreal, she created the International Laboratory for Brain, Music and Sound Research (BRAMS), a joint project of UM and McGill. Last fall, BRAMS received a total of $12 million in a matching grant from government and university sources. The money will allow the members of BRAMS, including Peretz, Zatorre, and nine other Montreal-based researchers, to explore music's mysteries. They seek to understand how humans cooperate to perform together, how children and adults learn to play music, and the relationship between music and language. "BRAMS will allow us to use music as a portal into the most complex aspects of human brain function," says Zatorre.

Already, BRAMS researchers have pinpointed areas of the brain involved in the perception of pitch. And they have demonstrated that pianists' ability to remember long, complicated pieces relies on close coordination between distinct motor and auditory memory circuits in the brain, in contrast to previous assumptions that the melody was key and that the fingers would simply follow.

Even before BRAMS was created, the Montreal group had established a world-class reputation in music research. "They are number one by a long shot," says Jamshed Bharucha, a music researcher and the provost of Tufts University in Medford, Massachusetts, who reviewed the BRAMS grant. Josef Rauschecker, a neuroscientist at Georgetown University Medical Center in Washington, D.C., also a grant reviewer, agrees. "The two of them, plus the colleagues they have assembled, are unbeatable as a team."

Music lover. BRAMS Co-Director Isabelle Peretz

The new laboratories on the UM campus include sound booths, a brain imaging lab, an echo-free chamber, and a performance lab sporting a $200,000 computerized Bosendorfer piano, which records the precise details of each keystroke and foot-pedal movement. One of BRAMS's most spectacular labs, however, will be located on the other side of town, at the concert hall of McGill's Schulich School of Music. Researchers plan to install wireless physiological sensors in up to 25 of the 188 seats to monitor group emotional reactions to live performances. The sensors will monitor heart rate, skin electrical responses, and even facial musculature of audience members who give informed consent, says McGill researcher Stephen McAdams, who heads this part of the project. Another 50 seats will be equipped with palm pilots, which subjects will be trained to use to indicate their reactions by moving a stylus "through emotional space," McAdams says.

These high-tech facilities will allow the BRAMS team to ramp up its research considerably. "We will be able to study musical perception and production in more natural contexts," says Peretz. "We can move from the study of single individuals to several subjects." The group is also designing musical instruments that can be played by subjects undergoing magnetic resonance imaging (MRI) brain scans. Zatorre says such studies will allow us to understand why different individuals have greater or lesser facility for learning, and what brain activity patterns differentiate a virtuoso pianist from a merely competent player."

For Peretz, 50, and Zatorre, 52, the creation of BRAMS is the high note in more than 25 years of research into music and the brain. But during much of that time, each of them was playing solo, and relatively few scientists were even listening. As a girl in Belgium, Peretz spent 11 years studying classical guitar. By the time she began university studies, however, "I knew that I was better at science." She received her Ph.D. at the Free University of Brussels in 1984 and came to McGill in 1985. With her doctoral supervisor Jose Morais, she proposed on theoretical grounds that the

CREDITS: (TOP TO BOTTOM) PAUL VANDENBERG/MCGILL UNIVERSITY; M. BALTE/SCIENCE

brain had specialized neural pathways for producing and responding to music, even if it shared some cognitive domains, such as hearing, with language. But many researchers were skeptical, arguing that the neural circuits involved in music mostly overlapped with those used in language and other functions.

At UM, Peretz began studying subjects suffering from congenital amusia, an inherited form of tone deafness that does not affect language or other mental functions. In the 1990s, she expanded this research to brain-damaged patients suffering from so-called acquired amusia, showing that many of them also had normal language abilities. That convinced many doubters that she was right. Music and language did seem to occupy different "modules" in the brain (*Science*, 1 June 2001, p. 1636). "Isabelle's work has fundamentally changed our understanding of how the brain processes music," says archaeologist Steven Mithen of the University of Reading in the United Kingdom, author of *The Singing Neanderthals*.

If there are special areas in the brain for music, where are they? That question attracted neuroimaging expert Zatorre, who, like Peretz, has had a lifelong affinity for music—he has played the organ since he was a child. Born in Buenos Aires, Argentina, Zatorre majored in both psychology and music at Boston University, where he played the organ at local churches, earning $25 for each service. He received his Ph.D. in experimental psychology at Brown University in 1981, and the same year started a postdoc at McGill, where he has been ever since. At McGill, Zatorre was in on the early days of the neuroimaging revolution, using MRI and other techniques to study the processing of both music and speech. His team was the first to demonstrate differences in brain activation patterns between musicians who have absolute pitch and those who don't, as well as the first to localize emotional responses to music in the brain. For example, Zatorre and colleagues showed that the "shivers-down-the-spine" feeling that many people get when they listen to pleasurable music correlates with activation of brain regions such as the amygdala and the orbitofrontal cortex, regions that are also associated with responses to food and sex.

Despite these accomplishments, until recently Peretz and Zatorre, like other music researchers, "were marginalized," Peretz says. Bharucha agrees: "I had an uphill battle convincing people that music was important as a cognitive or brain phenomenon." The turning point for Peretz and Zatorre came in 2000, when the New York Academy of Sciences asked the pair to organize a conference on the biological foundations of music. Soon afterward, the field began to explode. One reason, Zatorre says, was that brain researchers such as himself began collaborating with cognitive psychologists such as Peretz. "The cognition people started realizing, 'Hey, there is a brain out there.' And the neuroscientists and imaging people began realizing that they had something to contribute to the [cognitive] models."

The collaboration between Peretz and Zatorre, for example, has identified the brain region apparently defective in people with congenital amusia, which affects 5% or more of some populations. In the October 2006 issue of *Brain*, Peretz, Zatorre, and their co-workers showed that on MRI scans, tone-deaf individuals had less white matter in the right frontal inferior gyrus, an area just behind the right side of the forehead, compared to musically normal controls. The results were consistent with earlier work by Zatorre and others suggesting that the right frontal cortex is implicated in perceiving pitch and remembering music. In contrast, language is most often localized on the left side of the brain.

Yet whereas work on musical learning is showing progress, "music remains a mystery," Peretz says. "The biggest question is what it is for." Researchers have suggested many scenarios for why musical abilities might have evolved, such as enhancement of social solidarity and increasing communication between mothers and children (*Science*, 12 November 2004, p. 1120). But Peretz is cautious about such speculations. Although she has long argued against claims by researchers such as Harvard University cognitive scientist Steven Pinker that music is just "auditory cheesecake" with no adaptive function, she conceded in a recent review in *Cognition* that most hypotheses about music's role in human evolution are inherently untestable. "I believe that music is in our genes, but belief is not science"—more evidence is needed, she says.

Nevertheless, Zatorre says, "Pinker has served as a useful foil" for researchers who believe that music serves a unique biological role. And one reason for the explosion in this research, Zatorre adds, is that "music can serve as a probe into just about every mental function, including perception, motor performance, memory, attention, and emotion." In addition, some BRAMS research may have practical applications. For example, patients learning to speak again after strokes or other brain injuries do so more rapidly if they sing along with another person. Studying the neural changes that accompany such therapies could boost their effectiveness, says Peretz. Zatorre says he eventually wants to improve the technology of cochlear implants so the hard of hearing can listen to music. "Current implants are designed to maximize perception of speech," he says. "They are not good for music."

Musical mind. Robert Zatorre with computerized piano.

CREDIT: M. BALTER/SCIENCE

Although the creation of BRAMS seems likely to keep Peretz and Zatorre in the front ranks of music neuroscience research, both express some nostalgia for the musical careers they left behind. Zatorre is thrilled that the chapel in the former convent where BRAMS is housed has a working organ. "My dream is to do experiments in the morning and play in the afternoon," he says. Peretz still has her guitar. "I made my choice," she says, "but sometimes I still have regrets."

–MICHAEL BALTER

Still a believer Verfaillie in her Minnesota lab

STEM CELLS

Controversial Marrow Cells Coming Into Their Own?

Despite wide skepticism, Catherine Verfaillie persevered in her research and remains optimistic that her MAP cells will one day be useful in therapy

Catherine Verfaillie of the University of Minnesota made a big splash in 2002 when she reported in *Nature* that her lab had cultivated a type of cell with some seemingly remarkable properties. Called multipotent adult progenitor (MAP) cells and derived from the stromal cells of the bone marrow, the cells seemed to be able to turn into "most, if not all" cell types, the team wrote. And when injected into a mouse embryo, they appeared to contribute to most somatic cell types, creating a so-called chimeric mouse. MAP cells were surprisingly versatile for adult cells, whose fates were generally believed to be predetermined—in fact, they looked almost as good as embryonic stem (ES) cells.

Although Verfaillie was cautious in her claims, the press seized on MAP cells as an alternative to ES cells, which are controversial because creating them involves the destruction of an embryo. Scientists, too, were eager to try them out. But in the years since, MAP cells have failed to live up to expectations. They proved hard to grow, and the talk over coffee at meetings was that no one could replicate the *Nature* work.

Last month, however, Verfaillie's cells gained fresh attention. In the January issue of the *Journal of Experimental Medicine*, a group at Minnesota and Stanford University reported that they had used MAP cells to rebuild the blood system in mice. The work has impressed one skeptic. Stanford blood stem cell researcher Irving Weissman, who collaborated on the new work, Weissman calls the result "remarkable." His skepticism, he adds, "makes me a perfect collaborator because I insisted on very rigorous criteria for the experiments."

He emphasizes, as does Verfaillie, that these cells are clearly not as versatile as ES cells. But despite their limitations, they could prove to be useful therapeutically. "A lot of people have lost interest in MAP cells by this point," says Weissman. "What our paper will help do is get everybody to look at it again." Others agree. "I'm sure it will revive interest in MAP cells," says stem cell researcher Paul Schaffer of the University of Miami, Florida.

Going to the marrow

Bone marrow basically contains two types of stem cells: those that give rise to blood, and the stromal stem cells that develop into bone, fat, muscle, and cartilage. MAP cells are derived from early precursors in this latter population.

Verfaillie's team stumbled on MAP cells more or less by accident when, in trying to grow mesenchymal stem cells (a type within the heterogeneous stromal cell population), they came up with a culture system that seemed to select for even more primitive cells (*Science*, 21 June 2002, p. 2126). The tiny cells are strictly artifacts of lab culture, requiring at least 30 population doublings before displaying some of the characteristics of ES cells.

Verfaillie was not prepared for the extreme reaction when she went public with her findings in 2002. "Certainly the perception on the part of everyone was these cells were going to do it—that is, accomplish feats hitherto ascribed only to ES cells," says David Scadden, co-director of the Harvard Stem Cell Institute. But, he says, they turned out to be "a very fiddly cell to work with." Even scientists who obtained cells directly from Verfaillie couldn't make them perform. For example, Marcus Grompe, a liver stem cell researcher at Oregon Health & Science University in Portland, says he tried very hard but failed to get MAP cells to develop into hepatocytes in mice whose livers had been destroyed.

Verfaillie has retained her high hopes for MAP cells, although she acknowledges that in retrospect, the original chimera result might have been wrong. "We cannot exclude currently that that is not due to fusion" rather than chimerism, she says. Fusion, in which introduced cell types merge with the locals, first came to scientists' attention shortly before the Verfaillie *Nature* paper appeared (*Science*, 15 March 2002, p. 1989) and ultimately led to disillusionment with the notion that adult cells were capable of turning into other types of cells. Now researchers believe that in

CREDIT: UNIVERSITY OF MINNESOTA

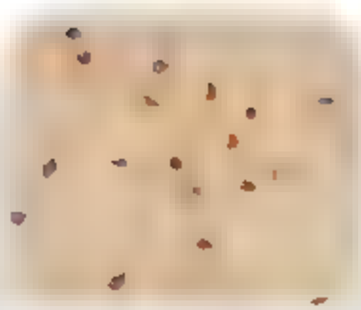
Stem Cell Candidates Proliferate

The stem cell landscape is getting crowded. Over the past few years, scientists have reported a variety of new types of stem cells, from both animal and human sources, including fetal liver, mouse testes, bone marrow, and umbilical cord blood. One of the most recent—and widely publicized—studies, by Anthony Atala of Wake Forest University in Winston Salem, North Carolina, described the potential of cells isolated from the amniotic fluid to differentiate into all three germ layers in vitro and into bone and brain cells in live mice (*Science*, 12 January, p. 170).

All the reports are preliminary, and none of the new cell types appears to be able to duplicate the potential of embryonic stem (ES) cells. Nonetheless, in combination, they may one day become prime players in stem cell therapy. A few examples:

• **USSCs:** A group at the University of Dusseldorf in Germany has identified a population of cells from human cord blood, which they call USSCs, or unrestricted somatic stem cells. They have "many overlapping features with MAP cells," they report, differentiating into a variety of tissues in vitro.

• **EPCs:** Douglas Losordo of Tufts University School of Medicine in Boston, Massachusetts, is currently conducting a study injecting bone marrow–derived cells called endothelial progenitor cells (EPCs) into patients with angina in hopes of creating new blood vessels to the heart.



Multiple progenitors. MAP cells, such as these from rat, are just one type of stem cell under investigation.

• **MIAMI cells:** A group at the University of Miami School of Medicine in Florida has reported a population of "pluripotent" cells from human bone marrow that they have dubbed marrow-isolated adult multilineage inducible (MIAMI) cells. MIAMI cells may complement MAP cells, they say, expressing the two stem cell markers, Nanog and Sox2, that MAP cells lack.

Robert Lanza of Advanced Cell Technology Inc. in Worcester, Massachusetts, predicts that a variety of different stem cells will prove optimal for different diseases. And in some cases, stroke for example, in which blood vessels and neurons are damaged, more than one stem cell type will be used. Marcus Grompe of the Oregon Health & Science University in Portland agrees that "there's a good chance" that new stem cell therapies will rely primarily on non-ES cells. ES cells are essential research tools. But, he says, at the current state of knowledge, "I haven't met a single person who wasn't leery of them for putting into people." —**C.H.**

many—perhaps most—reported cases of so-called adult cell plasticity, the appearance of cells switching identities was in reality misleading signals from fused cells.

But Verfaillie and Weissman ruled out fusion in the mouse blood paper and say the findings are indisputable. The researchers took mouse MAP cells and expanded them with up to 80 population doublings. They then transplanted them into 28 mice whose bone marrow had been destroyed by radiation. Because MAP cells are slow to grow and thus would be unable to repopulate the blood system fast enough to save a mouse from dying of radiation, they also injected blood stem cells. In the study, reported online 15 January in the *Journal of Experimental Medicine*, the MAP cells survived and contributed to the blood systems of 21 of the 28 mice.

Blood makers. MAP cells populate mouse lymph node.

Scientists must now understand that mouse MAP cells can make normal blood," proclaimed co-author Weissman. That makes them a promising treatment option because the population is readily expandable, in contrast to hematopoietic stem cells, which are difficult to expand in the lab. But first, he says, scientists would have to find a way to boost the speed at which MAP cells work, perhaps by pushing them to a more advanced stage of development before they are transplanted. And of course the study must be replicated with human MAP cells.

Verfaillie attributes their success in part to the development of improved culture conditions over the past half-dozen years, which are resulting in more homogeneous cell populations with high levels of Oct4, the main marker they share with ES cells. Verfaillie, who has been working at both Minnesota and the University of Leuven in Belgium for the past year and is now settled at Leuven as head of its stem cell institute, has also trained a number of groups, mainly in Europe, in MAP cell cultivation.

She says some 30 papers—about two-thirds with her as a co-author—have so far been published on MAP cells, and more are in the works. Last November, for example, Felipe Prosper's group at the University of Navarra in Pamplona, Spain, published a paper in *Blood* reporting that MAP cells contribute to tissues lining the walls of veins and arteries in mice.

Verfaillie also believes the cells may hold promise for mending sick livers and other organs. Some cells, she says, may "have to commit in the dish first—the environment in vivo may not give the right signals." She says a half-dozen papers currently in preparation or under review will present evidence that undifferentiated MAP cells can differentiate to specific cell types and be of therapeutic benefit in mice. Athersys, a biotech company in Cleveland, Ohio, is licensed to produce the cells, which are patented by the University of Minnesota.

Limitations

So far, says Verfaillie, what gives her group the most trouble is trying to make nerve cells and heart cells. Nerve cells are ectodermal tissue, and "it seems easier to make mesoderm and endoderm," she says. And even though heart tissue is mesodermal, no one has had any luck coaxing MAP cells to function as heart cells either. Even in vitro, "we still can't make cardiac anything, which is strange," says Verfaillie. Nonetheless, she's more optimistic about MAP cells' ultimate versatility than is Weissman, who remains skeptical despite the recent blood success.

It's still not clear exactly how MAP cells will stack up to ES cells. ES cells are the gold standard of pluripotency—which is usually defined as the capability of generating all cell types in the adult body. There are several important markers of pluripotency, namely Oct4, Nanog, and Sox2. The ability to form benign tumors called teratomas is one of the basic tests for pluripotency, as is the generation of a chimeric mouse from injecting a single cell into a mouse blastocyst. MAP cells have Oct 4, but they lack both Nanog and Sox2. Nor can they form teratomas. "We termed them 'multipotent' because the cells definitely do not have all the features of ES cells," says Verfaillie.

Verfaillie says she regrets the hype over her findings, which she sees as "in large part politically motivated"—and by no means confined to MAP cells. With these and other types of stem cells, she says, it is too soon to predict where treatments will be found.

—**CONSTANCE HOLDEN**

CREDITS: (TOP TO BOTTOM) CATHERINE VERFAILLIE AND ELIA PANOSALATYS-MONTAIN

LETTERS

edited by Etta Kavanagh

Limits to the Human Cancer Genome Project?

ALTHOUGH THE EFFORTS OF VOGELSTEIN AND COLLEAGUES TO DEFINE THE MUTATIONAL landscape of two of the most common human cancers (breast and colon) are highly commendable ("The consensus coding sequences of human breast and colorectal cancers," T. Sjöblom *et al.* Research Article, 13 Oct. 2006, p. 268), they also put into stark reality the challenges facing the Human Cancer Genome Project (HCGP). One wonders about the merits of such high-cost, low-efficiency, and ultimately descriptive-type "brute force" studies. Although previously unknown mutated genes were unearthed, the functional consequences of most of these and their actual role in tumorigenesis are unknown, and even with that knowledge we are a long way from identifying new therapeutic targets. Screening to identify potentially important genes in nonhematopoietic malignancies is now possible because of recent advances in transposon-based, unbiased, forward mutagenesis screens with potential for tissue-specific mutagenesis in mice (*1*, *2*). The advantage of such a system is that researchers can identify mutations that initiate, cooperate, and maintain a tumor by observing the development and progression of tumors in mice. Mouse models of human cancer are traditionally generated to contain known genetic changes identified in human cancers. Recent studies have demonstrated the power of the reverse, using abnormalities detected in cancer mouse models to study previously unknown, synthetic genetic lesions and their significance in human cancers (*3*, *4*). These comparative oncogenomic approaches combined with unbiased mutagenesis screens should provide a list of high-priority targets that can then be studied comprehensively for mutations and epigenetic abnormalities in human cancer and validated as therapeutic targets. Such approaches may be more rational and cost-effective, allowing a better compromise between achieving the major goals of the HCGP and appropriation of funding to other worthwhile cancer efforts.

WEE J. CHNG

Division of Hematology-Oncology, Mayo Clinic, Scottsdale, AZ 85259, USA. E-mail: chng.wee@mayo.edu

References

1. L. S. Collier, C. M. Carlson, S. Ravimohan, A. J. Dupuy, D. A. Largaespada, *Nature* **436**, 272 (2005).
2. A. J. Dupuy, K. Akagi, D. A. Largaespada, N. G. Copeland, N. A. Jenkins, *Nature* **436**, 221 (2005).
3. M. Kim *et al.*, *Cell* **125**, 1269 (2006).
4. L. Zender *et al.*, *Cell* **125**, 1253 (2006).

IN THE RESEARCH ARTICLE "THE CONSENsus coding sequences of human breast and colorectal cancers" (13 Oct. 2006, p. 268), T. Sjöblom *et al.* present the most extensive data available on sequencing of DNA from human cancers. Sequencing of 465 Mb from 22 tumors yielded 189 cancer-associated genes and reveals the enormous diversity and complexity of cancer genomes. The findings that "[t]he vast majority of genes were not known to be mutated in tumors," the average tumor harbored 90 mutated genes, no gene was consistently mutated in either breast or colorectal tumors, and there was no uniform panel of mutated genes argue against massive DNA sequencing being an efficient method for target identification. The authors' statement that "the number of mutational events occurring during the evolution of human tumors ... is much greater than previously thought" ignores a large literature on tumor heterogeneity and random mutations in human cancers (*1*).

There are important limitations to discovery of mutated genes by DNA sequencing. First, the vast majority of DNA sequence alterations are single-nucleotide polymorphisms, germline mutations, PCR errors, or DNA duplications. More than 99% of the nucleotide alterations observed by Sjöblom *et al.* were not somatic mutations. Second, conventional DNA sequencing does not detect nonclonal (random) mutations, and a consensus sequence is not informative of mutations in single cells. Although both random and clonal mutations could drive tumorigenesis, random mutations may also account for tumor cell heterogeneity, metastasis, and drug resistance (*2*). Third, expansion of the Human Cancer Genome Project must be rigorously justified in the context of diminished funding for the investigator-initiated grants that are critical for generating new approaches to improve cancer treatment and prognosis.

LAWRENCE A. LOEB AND JASON H. BIELAS

Department of Pathology, University of Washington, Seattle, WA 98195, USA.

References

1. L. A. Loeb, C. F. Springgate, N. Battula, *Cancer Res.* **34**, 2311 (1974).
2. J. H. Bielas *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18238 (2006).

IN THEIR RESEARCH ARTICLE "THE CONSENsus coding sequences of human breast and colorectal cancers" (13 Oct. 2006, p. 268), S. Sjöblom *et al.* report an initial total of 816,986 putative nucleotide changes in tumors. Of these, 33% or 256,957 silent changes were discarded as insignificant. A second screen uncovered 133,693 changes, but we are not told how many were silent

The winnowing process, which reduced the candidates to 191 changes, would probably have reduced the silent mutations proportionately. An independent screen of malignant gliomas after alkylation therapy found 26% silent mutations (*1*). Silent, or synonymous, mutations are assumed to be nonselective. A frequency of about 25% is expected if mutation is random. The frequency of silent mutations in the International Agency for Research on Cancer p53 database (*2*) is 4.41%, with a theoretical expectation of 23.8% for random mutation using the p53 selection of codons. Just as the low frequency of silent mutations in p53 indicates that this gene is selected during tumorigenesis, so their high frequency in this sequencing effort indicates that the vast majority of mutations occur randomly and are found in tumors as passengers. This conclusion in no way minimizes the importance of particular mutations for which the percentage of silent mutations is similar to that of p53. The Sanger Center Data Base, COSMIC, lists a total of 109 silent mutations out of 2335 mutations (4.67%) in eight cancer-related genes (e.g., CDKN2A, EGFR, KIT and RB1) for which more than 100 total mutations had been recorded. New cancer-related genes may possibly be identified by methods similar to those used by Sjoblom *et al*. The question is whether massive sequencing is an efficient way to uncover them against the huge background of random mutational noise.

BERNARD S. STRAUSS

Departments of Molecular Genetics and Cell Biology and Radiation and Cellular Oncology, University of Chicago, Chicago, IL 60637, USA.

References

1. C. Hunter *et al*., *Cancer Res.* **66**, 3987 (2006).
2. M. Olivier *et al*., *Hum. Mutat.* **19**, 607 (2002).

Response

MAGINE THAT NO MOLECULAR CANCER RESEARCH had ever been performed before the Sjoblom *et al*. study on colorectal cancers. In a single stroke, the study would have identified all of the consensus coding sequence genes now known to play a role in this tumor type, including *TP53*, *APC*, *KRAS*, *SMAD4*, *FBXW7*, *EPHA3*, *SMAD2* and *TGFBRII*. Many other genes not previously implicated in colorectal tumors were concomitantly discovered to be mutated at significant frequencies. These included genes that are central to the pathogenesis of other forms of cancer, such as *GNAS*, *VFL*, and *RET*. The study would also have provided unprecedented clues to the pathogenesis of inherited cancer syndromes such as familial adenomatous polyposis, neurofibromatosis, Li-Fraumeni syndrome, juvenile polyposis, and multiple endocrine neoplasia. Finally, it would have pointed to virtually all the major pathways currently known to be involved in neoplasia. Our study took less than a year to complete (once the technology was developed) and cost a tiny fraction of what was actually spent to discover a subset of the identified genes through conventional means. This effort demonstrated that unbiased genome sequencing is an extremely efficient way to discover cancer genes.

Chng, Strauss, and Lueb and Bielas raise a number of other points in their thought-provoking Letters, which we address here.

1) We have already pointed out that a large fraction of the mutations found in cancers were likely to be passenger mutations. Although the observed plethora of mutations may account for the clinical and biological heterogeneity of tumors, most of them are not likely to be integral to neoplasia. Precisely for this reason, we designed a measure (CaMP scores) to rank genes via their mutation frequency, taking into account gene size and nucleotide context. Those genes with the highest CaMP scores are the ones of most interest for future genetic and functional studies. All of the genes noted above ranked among the top 30 such genes.

2) Our data strongly support previous results indicating that most human cancers have a mutation rate that is similar to that observed in normal cells (*1*). A small fraction of cancers have higher mutation rates because of mismatch repair deficiencies. Such cancers were excluded from our study because mutations in these tumors are more difficult to interpret.

3) Our study was designed to identify those mutations that may drive tumorigenesis, i.e., clonal mutations. Such mutations have been shown to be directly responsible for tumor progression and are the only ones known to be useful as diagnostic and therapeutic targets. Once clonal mutations are identified in late-stage or therapy-resistant tumors, it could be useful to search for such alterations in earlier stage clinical samples wherein they may be present in a small fraction of tumor cells. However, genome-wide identification of alterations present in small tumor subpopulations is neither feasible nor desirable at this time.

4) Sequencing studies cannot replace functional studies in model organisms, and the latter studies are essential to reduce morbidity and mortality from cancer. But sequencing studies *can* guide functional studies by focusing them on genes that are likely to play a role in human cancers. We have demonstrated that such genes can be identified by relatively simple and inexpensive sequencing methods. As technology improves, such sequencing will become even more cost-effective. To turn the process on its head by trying to identify genes important in human cancer through functional studies, followed only later by sequencing, would be substantially more costly and less comprehensive.

TOBIAS SJÖBLOM,[1] SIÂN JONES,[2] LAURA D. WOOD,[1] D. WILLIAMS PARSONS,[3] JIMMY LIN,[1] THOMAS BARBER,[1*] DIANA MANDELKER,[1] REBECCA J. LEARY,[1] JANINE PTAK,[1] NATALIE SILLIMAN,[1] STEVE SZABO,[1] PHILLIP BUCKHAULTS,[2] CHRISTOPHER FARRELL,[2] PAUL MEEH,[4] SANFORD D. MARKOWITZ,[5] JOSEPH WILLIS,[6] DAWN DAWSON,[6] JAMES K. V. WILLSON,[5] ADI F. GAZDAR,[6] JAMES HARTIGAN,[7] LEO WU,[8] CHANGSHENG LIU,[8] GIOVANNI PARMIGIANI,[9] BEN HO PARK,[10] KURTIS E. BACHMAN,[11] NICKOLAS PAPADOPOULOS,[1] BERT VOGELSTEIN,[1] KENNETH W. KINZLER,[1] VICTOR E. VELCULESCU[1]

[1]Ludwig Center and Howard Hughes Medical Institute, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD 21231, USA. [2]Department of Pathology and Microbiology, Center for Colon Cancer Research, and South Carolina Cancer Center, Division of Basic Research, University of South Carolina School of Medicine, Columbia, SC 29229, USA. [3]Department of Medicine, Ireland Cancer Center, and Howard Hughes Medical Institute, Case Western Reserve University and University Hospitals of Cleveland, Cleveland, OH 44106, USA. [4]Department of Pathology and Ireland Cancer Center, Case Western Reserve University and University Hospitals of Cleveland, Cleveland, OH 44106, USA. [5]Harold C. Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. [6]Hamon Center for Therapeutic Oncology Research and Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. [7]Agencourt Bioscience Corporation, Beverly, MA 01915, USA. [8]Soft Genetics LLC, State College, PA 16803, USA. [9]Departments of Oncology, Biostatistics, and Pathology, Johns Hopkins Medical Institutions, Baltimore, MD 21205, USA. [10]Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, MD 21231, USA. [11]Department of Radiation Oncology and Department of Biochemistry and Molecular

Biology, Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, MD 21201, USA.

*Present address: Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN 46285, USA.

Reference

1. C. Lengauer, K. W. Kinzler, B. Vogelstein, *Nature* **396**, 643 (1998).

Gene Expression and Ethnic Differences

IN HER ARTICLE "IN ASIANS AND WHITES, GENE expression varies by race" (News of the Week, 12 Jan., p. 173), J. Couzin explains that recent research has demonstrated that there is significant variation in gene expression in Asians and whites, although providing the disclaimer that "[g]enetic variation among races, long a political hot potato, has also been a scientific puzzle." Indeed, anthropologists and geneticists alike have been keenly interested in determining just what makes us different at the genetic level. The problem is how these differences are conceptualized. After decades of discussion, we are still left asking "what is a race?" Race is truly a biological and taxonomic problem, not simply a sociological problem as many have argued. Now, it should be asked, what does it mean to say that gene expression varies between "Asians" and "whites." In the context of this research, "Asian" means either Japanese or Chinese or both groups pooled together. The term "white" is used interchangeably with "Caucasian" and people of "European descent." This imprecise use of language further adds to the problem of not being able to accurately describe or classify the reference population. Who do these people represent? Certainly, both Asia and Europe are very large and diverse continents, both socially and biologically. The difference in gene expression between the Japanese and Chinese subjects cited in this report demonstrates that variation occurs not only on a continental level but also locally, which is indicative of patterns of ethnic variation rather than racial variation. We should not simply be asking how human groups vary in terms of genetic composition, but how those groups vary in terms of the social and biological processes that created them.

MICHAEL S. BILLINGER

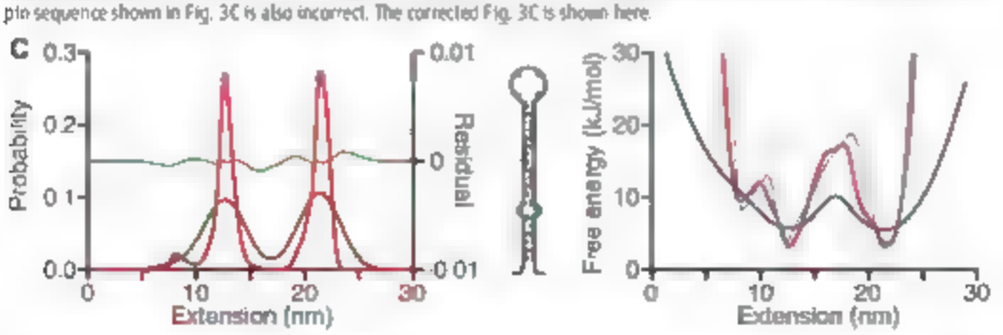
Edmonton, Canada.

CORRECTIONS AND CLARIFICATIONS

Brevia: "Ancient noncoding elements conserved in the human genome" by B. Venkatesh *et al.* (22 Dec. 2006, p. 1892). Table S1 should have been included in the Supporting Online Material (SOM) rather than posted on the author's Web site. Table S1 includes the coordinates of the conserved noncoding sequences within the elephant shark genome and the corresponding coordinates in the human genome. The elephant shark sequences (1.4× coverage) have been deposited at DDBJ/EMBL/GenBank under the project accession AAVX00000000. The version described in this paper is the first version, AAVX01000000. The authors have also submitted the traces of these sequences to the Trace Archive at the NCBI. Tables S8 and S9 should also have been included in the SOM, and the following statement should have been added to the acknowledgments in reference 8: "The sequences of the human and elephant shark noncoding elements are in Tables S8 and S9. Tables S1, S8, and S9 are now available with the SOM at www.sciencemag.org/cgi/content/full/314/5807/1892/DC1."

Research Articles: "The genome of the sea urchin *Strongylocentrotus purpuratus*" by Sea Urchin Genome Sequencing Consortium (10 Nov. 2006, p. 941). On pages 951 and 952, errors were made in renumbering authors' affiliations. Some changes were missed, and the affiliation for N. Nikki Adams was omitted. C. G. Elsik, I. Hrbinko, and V. D. Vacquier appear twice. C. G. Elsik is at Texas A&M University. T. Hrbinko is at the Sunnybrook Research Institute and Department of Medical Biophysics, University of Toronto. V. D. Vacquier is at the Scripps Institution of Oceanography. Corrected group affiliations, then individuals alphabetically: P. Kitts, M. J. Landrum, D. Maglott, K. Pruitt, A. Souvorov, National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD 20894, USA. O. Fedrigo, A. Primus, R. Satija, Department of Biology and Institute for Genome Sciences and Policy, Duke University, Durham, NC 27708, USA. Nikki Adams, Biology Department, California Polytechnic State University, San Luis Obispo, CA 93407, USA. C. Flytzanis, Department of Biology, University of Patras, Patras, Greece, and the Department of Molecular and Cellular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. B. E. Galindo, Biotechnology Institute, Universidad Nacional Autónoma de Mexico (UNAM), Cuernavaca, Morelos, Mexico 62250. J. V. Goldstone, Department of Molecular, Cellular, and Developmental Biology, University of California, Berkeley, Berkeley, CA 94720, USA. G. Manning, Razavi-Newman Center for Bioinformatics, Salk Institute for Biological Studies, La Jolla, CA 92186, USA. D. Meritt, Center of Marine Biotechnology, University of Maryland Biotechnology Institute, Columbus Center, Baltimore, MD 21202, USA. J. Song, Department of Molecular and Cellular Biology and Biochemistry, Brown University, Providence, RI 02912, USA. D. P. Terrilliger, Department of Biological Sciences, George Washington University, Washington, DC 20052, USA. A. Wikramanayake, Department of Zoology, University of Hawai'i at Manoa, Honolulu, HI 96822, USA.

Reports: "Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid" by M. T. Woodside *et al.* (10 Nov. 2006, p. 1001). On page 1002, the key in Fig. 2C is incorrect. In the legend for Fig. 2C, the descriptions of the color blocks should read, "Distance from F to U (black), F to F (blue), and U to U (red) plotted versus the mismatch location." The hairpin sequence shown in Fig. 3C is also incorrect. The corrected Fig. 3C is shown here.



News Focus: "Truth and consequences" by J. Couzin (1 Sept. 2006, p. 1222). Due to an editing error, the sentence on page 1223, "Koersten and Padilla talked for about an hour and together examined the papers cited in the proposal" was incorrect. It should have read, "Koersten and Padilla talked for about an hour and together examined the pages of the proposal."

TECHNICAL COMMENT ABSTRACTS

COMMENT ON "Obestatin, a Peptide Encoded by the Ghrelin Gene, Opposes Ghrelin's Effects on Food Intake"

N. Chartrel, R. Alvear-Perez, J. Leprince, K. Iturrioz, A. Reaux-Le Goazigo, V. Audinot, P. Chomarat, F. Coge, O. Nosjean, M. Rodriguez, J. P. Galizzi, J. A. Boutin, H. Vaudry, C. Llorens-Cortes

Zhang *et al.* (Research Articles, 11 November 2005, p. 996) reported that obestatin, a peptide derived from the ghrelin precursor, activated the orphan G protein–coupled receptor GPR39. However, we found that ^{125}I-obestatin does not bind GPR39 and observed no effects of obestatin on GPR39 transfected cells in various functional assays (cyclic adenosine monophosphate production, calcium mobilization, and GPR39 internalization). Our results indicate that obestatin is not the cognate ligand for GPR39.

Full text at www.sciencemag.org/cgi/content/full/315/5813/766c

RESPONSE TO COMMENT ON "Obestatin, a Peptide Encoded by the Ghrelin Gene, Opposes Ghrelin's Effects on Food Intake"

Jian V. Zhang, Cynthia Klein, Per-Gen Ren, Stefan Kass, Luc Ver Donck, Dieder Moechars, Aaron J. W. Hsueh

We cannot reproduce our original findings on obestatin binding and activation of GPR39 receptors in vitro. However, we can reproduce our original findings on the in vivo effects of obestatin in mice (decreases in food intake, gastric emptying responses, and body weight gain) under precise experimental conditions. Further studies are needed to reveal the exact relation between obestatin and the G protein–coupled receptor GPR39.

Full text at www.sciencemag.org/cgi/content/full/315/5813/766d

POETRY

Metaphors Be with You

Diane Ackerman

All freshmen entering Boston University in 1966 were assigned C. P. Snow's *The Two Cultures* (*1*) to read—a fitting choice given the reality of academic life at U.S. colleges, where students (however wide-ranging and combinable their interests) had to choose between the sciences and the humanities. Fascinated by psychology and the flourishing field of neuroscience, I intended to major in bio-psychology. But when I transferred colleges a year later, a computer mistakenly placed me in English. Because I had been writing poetry shyly but enthusiastically all my life, I considered it late.

Contemporary Poetry and Contemporary Science
Robert Crawford, Ed.
Oxford University Press, Oxford, 2006. 288 pp. $35, £19.99. ISBN 9780199258178

By the time I started graduate work at Cornell, little had changed. Students still had to choose between poetry and science, and, if anything, the divide seemed larger. Physicist Richard Feynman declared that the universe could only be understood in the language of mathematics, an argot I didn't speak. The literary critics I met regarded science as antithetical to both art and genuine feeling. A renowned scholar took me aside one day and said, "What's a nice girl like you doing writing about amino acids?"

I had been listening to Gustav Holst's *The Planets*, enchanted by its lyrical flights, but also puzzled by our apparent need to imagine Venus as a goddess of love or Mars as the bringer of war in order to find them captivating, when the physical reality of the planets offered artists and scientists alike startling new views of what we summarily call nature: a cornucopia of picturesque landscapes, fresh metaphors, and elemental novelties ripe for wordplay as well as a gradually widening aperture of belief in what surprises still lay hidden from view just over the next fence post, underfoot, or in orbit. A great fan of the universe, which I took literally as one verse, I decided to write a suite of scientifically accurate poems about the planets. And, since I'd been reading Metaphysical, Imagist, and ancient Greek poets, among others, who embraced the revelations of science in their work, I chose a Ph.D. dissertation on the metaphysical mind in science and poetry, as one of the fascinating ways that the mind works.

That was in 1972, the last time when it was possible to learn everything known about the planets. (The best NASA photo of Pluto was a tiny ball of light with an arrow pointing to it.) Carl Sagan agreed to be my technical adviser on *The Planets: A Cosmic Pastoral*, as well as a member of my doctoral committee, which also included the poet A. R. Ammons. With that duo running interference for me, I was able to pursue studies in poetry and science, pore over NASA photos and maps, attend planetary fly-bys, and audit lectures in the so-called sciences (I regarded it all as nature) while earning degrees in creative writing and English literature.

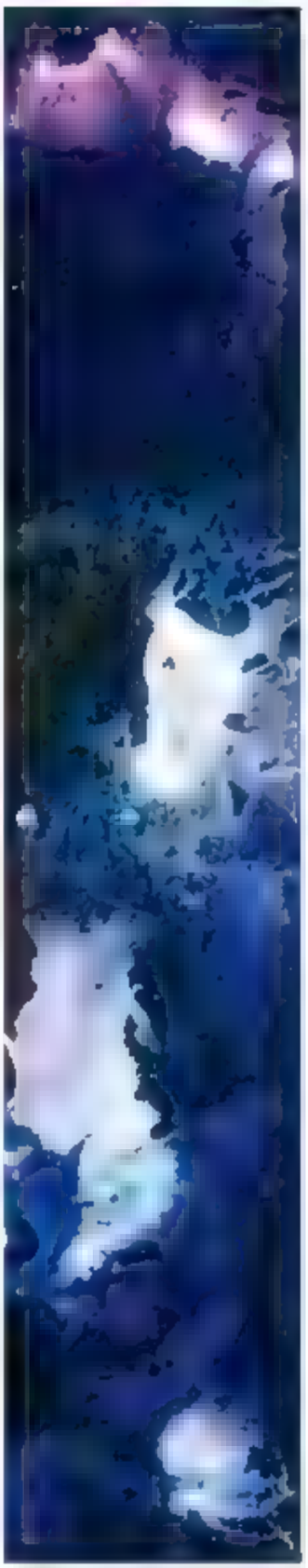
Such memories arrived unbidden as I read *Contemporary Poetry and Contemporary Science*. The volume is an offspring of a project that began with a group of poets at the 1998 Edinburgh International Science Festival and led to poet-scientist encounters, often over lunch or in a lab, with the goal of a "meaningful interchange between poetry and science."

Is that interchange really still a frontier? On many campuses, yes—in part because departments have become *matryoshka* dolls of subspecialties, where one usually prospers by narrowing the view. But less so, I think, in society in general, where scientists and humanists often meet at the crossroads of necessity to discuss neuroethics, contemplative neuroscience, the sense and sensibility of animals, or stem cell research, among other topics. Popular science television programs, science cafés, and interdisciplinary books have all helped to bridge the cavernous mélange.

This most recent offering begins with immunologist and poet Miroslav Holub distinguishing the rational passion of science from the emotional bogs and dreamy mindscapes of poetry. Holub is "astonished by the frequent incidence of failure at high school (or failure at university) that features in the biographies of poets." He insists that "a scientist (even writing poetry) should be and must remain a scientist. He is the member of the intellectual community who still has a deep and visible importance for the society and its everyday life." That's bound to ruffle some feathers.

Poet Edwin Morgan offers an antidote: a smart, far-ranging essay that touches on Lucretius' phenomena, eye, Goethe's study of optics, Shelley's giving himself electric shocks to see if his hair would stand on end (it did), Leopardi pawing distant galaxies, and Philip Larkin once declaring that he couldn't "stand anything to do with the wonders of nature," among many other topics. In an imaginary dialogue (originally written in 1957) between neurophysiologist Grey Walter and filmmaker Jean Cocteau, Morgan has Cocteau describe poetry as "an act of love with an angel in anger … a darkness for the searchlight of a question." Walter's reply includes the observation that "Machines can be persuaded to stumble on dreams—Except that it isn't stumbling and they aren't dreams."

Other engaging essays follow, studded with wit and humor. Because they benefit from many digressions, it's hard to epitomize them.



CREDIT: COURTESY SANDI RITCHIE MILLER

Sandi Ritchie Miller's *Ocean of Space*.

The reviewer is the author of *An Alchemy of Mind: The Marvel and Mystery of the Brain*. Web site: www.dianeackerman.com

but I'll choose a few salience to suggest their range of thought. Poet W. N. Herbert asks, "Can poets write meaningfully about science, or are they instead limited to writing poetry that employs scientific terminology?" Poet John Burnside champions the "beauty of the real ... The starlit darkness of the actual night, the salt and physicality and achieved grace of real bodies." Without "necessary awe," he writes, any description of the world is a lie. Poet Simon Armitage defines science as "an ongoing refinement of metaphor. There is no such thing as a molecular structure, only a little model of Ping-Pong balls held together by pipe-cleaners to help us believe in it." Literary critic Adelaide Morris considers both science and poetry as thought experiments and reasons a la Einstein that "[i]t is the theory which decides what we can observe." Astronomer Jocelyn Bell Burnell writes with elegant enthusiasm about poems recruiting space imagery and how poetry heals and enriches the many strata of her life. Reflecting on her collection of "120 or so astronomy poems" written since 1950, she wonders why poets seem to avoid the new wavelengths (few refer to radio telescopes, none to x-ray astronomy). Her essay provides a superb guided tour of astronomy poems, pausing insightfully at many ports of call.

Psychiatrist Kay Redfield Jamison asks, "Can a writer's life and work really be understood without at least a rudimentary scientific understanding of emotions, temperament, cognitive psychology, and genetics?" Realizing that Milton, for example, could write "All Hell broke loose" because he knew exactly where Hell lay, having sent his wife and daughters there often enough, might just bias one's reading of *Paradise Lost*, which would be a shame. And yet, how dare we ignore the underlying moods? This question has fueled many polite (and impolite) literary quarrels. How does one savor Bruno Schulz's phantasmagorical prose and Chagall-like paintings, for example, and completely ignore the dominatrix obsessions of his artworks?

In all, 21 essays by scientists, poets, and critics and a handful of poems fill out this entertaining and illuminating volume. As the stepchild of brief encounters, *Contemporary Poetry and Contemporary Science*, refreshingly, doesn't aim for coherence or even decorum, but offers contingency samples from a frontier that, happily, is much more hospitable than it used to be. When cybernetician Kevin Warwick (who once had a silicon chip transponder implanted in his arm) read poet Michael Donaghy's poem "Grimoire," he felt that "Through Michael's words I was able to look at myself in a ten-dimensional space." That's a wide berth, fit equally for string theory or ballade.

Reference

1. C. P. Snow, *The Two Cultures and the Scientific Revolution* (Cambridge Univ. Press, Cambridge, 1959).

10.1126/science.1138360

HISTORY OF SCIENCE

Franklin's Civic Science

H. Frederick Dylla

In last year's *Rising Above the Gathering Storm* (*1*), the National Academies set forth a strategy for ensuring "that the United States is the premier place in the world for innovation." I often think that an important component of success is to have a clear vision of American science's past. Another book from 2006 offers a look back that may help enable such a look forward: Joyce E. Chaplin's *The First Scientific American: Benjamin Franklin and the Pursuit of Genius*.

Franklin "is currently celebrated as an American statesman or a founder, a politician who happened to do a little science on the side," writes Chaplin, a Harvard University historian of science. Then she notes something of interest in our current age when science and the rest of society often seem poles apart: "But that description puts the cart before the horse. Franklin was not a statesman who did science. He became a statesman because he had done science." And he was able to do so because, in the eighteenth century, science became part of public culture.

"Experiments and demonstrations were public events," Chaplin comments. The term scientist was not coined until the 19th century, and many citizens from all professions had a hand in advancing and appreciating scientific knowledge. For his famous electrical experiments, Franklin collaborated with a lawyer (Thomas Hopkinson), a silversmith (Philip Syng), and a Baptist minister (Ebenezer Kinnersley). In contrast to the polarization too often seen nowadays, how refreshing to think of "[m]embers of the clergy preach[ing] the new philosophy of nature."

Franklin's fame as a scientist comes from his legendary demonstration that lightning is electrical and from experiments that led to the principle of charge conservation. Chaplin's book, however, presents a many-decades-long example of something that might be just as important: how science and the rest of culture once interpenetrated deeply.

Chaplin also illuminates the whole spectrum of Franklin's scientific and technological achievements and activities. Her account demonstrates why Franklin became an international celebrity. The book delves into work that often goes unmentioned in popularized stories about Franklin and receives only slight attention in other biographies. In addition to covering Franklin's investigations of electricity, Chaplin offers excellent accounts of, for example, his interest in and studies of ocean currents and navigation (as in his description of the Gulf Stream) and his observation of the dispersion of oil on water (which amounted to an early entry into the field of monolayer films). She also documents Franklin's extensive considerations of heat flow, which have entered popular memory through the invention of the "Franklin" stove.

Chaplin chose excellent figures that illuminate Franklin's scientific contributions—such as his maps of the Gulf Stream, an assembly diagram for the Franklin stove, and his musical instrument (the "glass armonica")—as well as the evolution of portraiture depicting him as a man of science. Unfortunately, many of the illustrations are poorly reproduced, and some are nearly illegible.

Anyone, including any scientist, who writes a book as wide-ranging as this one risks the occasional technical or scientific error, and Chaplin is no exception. But the few, minor inaccuracies in the book do not distract from her elegant demonstration of the unity of Franklin's scientific and civic activities or from that demonstration's ability to provoke thoughts about future science. Moreover, Chaplin's diverse, eclectic, extensive mix of research sources suggests many years' worth of useful and enjoyable further reading.

Reference

1. National Research Council, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future* (National Academies Press, Washington, DC, 2006). www.nap.edu/catalog/11463.html.

The reviewer is at the American Institute of Physics, One Physics Ellipse, College Park, MD 20740–3843, USA. E-mail: dylla@aip.org

10.1126/science.1138024

SCIENCE AND SOCIETY

Time for a New Era of Science Diplomacy

Scientists are among America's most effective diplomats

Kristin M. Lord[1] and Vaughan C. Turekian[2]*

Science diplomacy has played an important if underappreciated role in U.S. foreign policy over the past 50 years. During the Cold War, the development of organizations such as the International Institute for Applied Systems Analysts and scientific exchanges between American and Soviet scientists provided a critical connection between adversaries (1). More than 30 years ago, robust scientific exchanges between the United States and China laid the groundwork for a relationship that has grown increasingly deep and complex (2). In recent years, scientific engagement has been a clear signal of friendship between the United States and countries such as India, Egypt, and Pakistan with high-level participation from the diplomatic communities in both countries (3) (see figure, right).

It is time to adopt science diplomacy for a new era. Old-fashioned diplomacy between governments, while necessary, is no longer sufficient. In this age of the Internet, rapid and relatively low-cost travel, and 24-hour global news, the power of nongovernmental organizations (NGOs), private companies, and social networks is rising. To protect and advance U.S. interests, the U.S. government needs to accelerate its engagement with these new actors and to build positive relationships with foreign publics, as well as their diplomats.

Science and technology (S&T) offer a promising entry point for engaging citizens and civil society organizations worldwide. An opinion piece on the op-ed page of the *Washington Times* called on the U.S. government to recognize the opportunities afforded by widespread respect for American S&T (4). We concur. But without the engagement and commitment of the U.S. scientific community, the government cannot succeed.

S&T make good diplomacy. Secretary of State Condoleezza Rice exchanges documents with India's Minister of Science and Technology, Kapil Sibal, after signing an umbrella S&T agreement, 17 October 2006.

Why Diplomacy Through Science?

Nearly 4 years ago, the United States entered a complex and difficult war with Iraq. Since then, global public opinion regarding the United States has reached all-time lows. Polls in 33 countries indicate that only 40% of those surveyed view America's influence in the world as mainly positive. In contrast, 45% view China positively and 58% hold favorable views of Europe. Dislike of America extends to long-time friends and allies. Only 30% of Canadians, 21% of Germans, and 15% of Turks hold favorable opinions of the United States (5).

Readers may ask why this matters if the U.S. government is charged with protecting U.S. interests, not winning popularity contests overseas. The answer is that increasingly, our interests depend on the support (or at least acquiescence) of foreign populations. Negative images of the United States translate directly into constraints on American influence and ability to implement policy.

Engaging foreign citizens is the goal of public diplomacy. As the 9/11 Commission report (6) underscored, engaging foreign public opinion is vital to winning the global struggle of ideas. President Bush clearly agrees, having appointed one of his closest advisers, Karen Hughes, to the position of Under Secretary of State for Public Diplomacy and Public Affairs in 2005. Despite this high-level attention, however, promoting America's image overseas continues to be a daunting task, particularly in predominantly Muslim countries.

In our view, public diplomacy is most effective when exercised through deeds rather than words. The U.S. government should focus on doing things that positively affect foreign societies and speak to what we stand for as a nation. We should foster tangible initiatives that promote education, economic growth, human well-being, and hope.

If we understand public diplomacy in these terms, the role of S&T is pivotal. Scientific education creates citizens with the critical thinking skills necessary for successful participatory governance and competition in the global economy. S&T are linked strongly with economic development (7). Zogby public opinion polls in several Middle Eastern nations, where the United States is particularly unpopular, indicate that S&T are the single most respected elements of American society. (8). Social science research indicates that collaboration to solve common problems is one of the best ways to foster positive relations between groups (9).

Under Secretary Hughes increasingly appreciates the unifying power of science (10). She held major events on breast cancer awareness in the Middle East and has launched a project on fighting malaria (11). She has created new International Science Fulbright awards and supported activities, led by Under Secretary of State for Democracy and Global Affairs, Paula Dobriansky, to cultivate networks of women scientists and engineers (12).

Yet in an era where international skepticism about U.S. foreign policy abounds, government can only do so much. Ultimately, civil society—including scientists and engineers—will need to join in this diplomacy of deeds in order for the new science diplomacy to succeed. The fact that science is, and should remain, outside the realm of politics only makes scientists better suited for this task.

How can the U.S. science community contribute to science diplomacy and remind the world that Americans are defined by more than specific U.S. government policies? Individual scientists can contribute by realizing that they are valuable ambassadors of goodwill. They can intensify their global activities and promote greater engagement with counterparts world-

[1]Elliott School of International Affairs, The George Washington University, Washington, DC 20052, USA. [2]Chief International Officer, AAAS, Washington, DC 20005, USA.

*Author for correspondence: vturekia@aaas.org

CREDIT: KEVIN WOLF/AP PHOTO

wide. They can increase their efforts to invite foreign peers to review scientific articles and papers. Senior U.S. researchers can use their own international networks, including former students and postdocs working outside of the United States, to reach out to junior scientists in other countries, to collaborate with peers, and to promote broader international cooperation. U.S. scientists should make a special effort to engage with scientists from countries where the United States is misunderstood or disliked—not to justify or promote any government policy, but to build bridges and trust. They can engage more with university students and the general public overseas, not just other scientists, and let them know how scientists from all nations make a collective difference in their lives. In so doing, U.S. scientists will make the world a better place and perhaps improve foreigners' views of America along the way.

International Conference on Women Leaders in Science, Technology, and Engineering, Kuwait City, Kuwait, 8 to 10 January 2007.

Scientists can also encourage their universities, research institutions, professional societies, and laboratories to adopt global engagement as a priority. Although a large sum of individual efforts is important, effective global engagement will be most influential if it engages whole organizations as well. Many of the major U.S. scientific and engineering societies already have specific offices or initiatives dedicated to international collaborations. To give just one example, in January AAAS joined the U.S. Department of State, the Kuwaiti government, and a Kuwaiti science NGO to convene a conference in Kuwait City to promote networks of women scientists and engineers in the broader Middle East (see figure above). AAAS has also recently started a pilot program that remotely links U.S. researchers with university-level science students in developing and emerging countries in order to share and discuss seminal papers across a range of scientific disciplines (*13*). Yet, despite current efforts, scientific organizations can do more.

Of course, all this assumes that scientists and engineers are willing to be ambassadors and to participate in the new science diplomacy. Why would they? The answer is threefold. First, while science holds great benefits for diplomacy, diplomacy also benefits science. For instance, in large-scale programs such as International Thermonuclear Experimental Reactor (ITER) (*14*), scientists from major powers such as China, India, Russia, Korea, Japan, the European Union, and the United States will work together in an unprecedented international agreement to develop fusion energy. Moreover, diplomacy can create opportunities to conduct research in parts of the world critical to scientific advancement. Scientific research ranging from astronomical observation in Australia to archaeological research in Libya depends on broader access, as well as diplomatic support.

Second, the health of the U.S. scientific community depends on the continued willingness of foreign scientists and students to come to the United States for study, research, and work. Visa difficulties, combined with a perception that the United States does not welcome foreigners, reduced the number of foreign students coming to the United States after 9-11. This trend is beginning to reverse, but negative perceptions persist and it is important to remain vigilant. The U.S. economy benefits greatly from foreign scientists and science (*15*). We must ensure that the United States remains attractive and welcoming.

Third, scientists are citizens. Like their counterparts outside of the scientific community, many scientists and engineers share concerns about negative perceptions of the United States. The good news is that scientists have some ability to change those perceptions for the better.

The Way Forward

Who should lead a renewed effort toward science diplomacy? Unfortunately, there is currently no ideal U.S. government agency to lead a sustained effort. Technical agencies, such as the Department of Energy, National Institutes of Health, and National Aeronautics and Space Administration are (not inappropriately) focused on their core missions and interested in international collaboration to the extent it advances those missions. The National Science Foundation has a broader mandate than these agencies, but NSF's goal is to foster basic research. Consequently, its international activities are designed to address specific research questions. The Department of State is designed to focus on diplomacy—but, unfortunately, is not well equipped to engage in science diplomacy. With limited resources for S&T cooperation, limited scientific expertise, and pressure to focus on the day's crises rather than long-term engagement, the department's efforts must be complemented by the work of other agencies. Unless those resources increase dramatically (which we do not believe is likely), the Department of State will need much more support in the area of science diplomacy.

It is time for the scientific community to increase its role in diplomacy—and maybe even take the lead. Nongovernmental scientific organizations are more credible, more nimble, and—as honest brokers—in many cases more respected than the U.S. government overseas. They work at the grassroots level on global problems such as energy, clean water, and health. A vigorous new science diplomacy, oriented to foreign citizens as well as their governments, will promote human well-being, will benefit science, and will catalyze public diplomacy. Our country needs a new era of science diplomacy, and we need the commitment of the U.S. science community behind it.

References and Notes

1. Further information about the founding of the International Institute for Applied Systems Analysis can be found at www.iiasa.ac.at/.
2. Z. Wang, *Hist. Stud. Phys. Biol. Sci.* **30** (part 1), 249 (1999).
3. For more information about U.S. bilateral science agreements, see www.state.gov/g/oes/stc/.
4. M. B. d'Arcy, M. A. Levi, *Washington Times*, 17 August 2005.
5. BBC World Service poll, 2006, www.worldpublicopinion.org/pipa/articles/views_on_countries/region_bv168.php?nid=&id=&pnt=168.
6. See www.9-11commission.gov/report/index.htm.
7. National Research Council, *The Fundamental Role of Science and Technology in International Development: An Imperative for the U.S. Agency for International Development* (National Academies Press, Washington, DC, 2006).
8. J. Zogby, J. Zogby, *Impressions of America: How Arabs View America, How Arabs Learn About America* (Zogby International, Washington, DC, 2004).
9. For an overview, see K. M. Lord, "What academics (should have to) say about public diplomacy," presented at the American Political Science Association Political Communication conference, Washington, DC, 31 August 2005.
10. In her 5 January 2007 interview with the *Chronicle on Higher Education*, Under Secretary Hughes discusses different methods for using science and education as an important tool for diplomacy: http://chronicle.com/weekly/v53/i18/18a03501.htm.
11. For a description of the Breast Cancer Awareness program, see www.state.gov/documents/organization/74175.pdf.
12. For an announcement of the new Fulbright awards, see www.state.gov/r/pa/prs/ps/2006/76482.htm.
13. Developing and Emerging Country Outreach Program, http://sustsci.aaas.org/edev/edev_survey.html.
14. ITER, www.iter.org/.
15. National Academy of Science, National Academy of Engineering, and Institute of Medicine, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future* (National Academies Press, Washington, DC, 2005).

10.1126/science.1139880

CREDIT: VAUGHAN TIBEDAN, AAAS

BIOPHYSICS

Antennae as Gyroscopes

Moths use their antennae as part of a motion sensing system that ensures stable flight

R. McNeill Alexander

Flying insects need to detect unwanted movements of their own bodies, so that they can make any necessary corrections to restore the status quo. They need to know, for example, when their flight is disturbed by an eddy in turbulent air or by an imperfectly executed wing beat. Dragonflies depend on sight for this information. That works well in bright daylight but would not be satisfactory in near-darkness because eyes cannot provide precise information quickly in dim light. Moths active at night need information about unwanted movements to maintain flight stability, especially when hovering to collect nectar from flowers. On page 863 of this issue, Sane and colleagues (*1*) explain how a hawk moth senses its own rotations.

These researchers found that the moth's movement-detection system depends largely on the Coriolis effect, which keeps spinning gyroscopes stable. This effect is an apparent deflection of an object viewed in a rotating frame of reference, seemingly attributable to an apparent force. We already knew of the importance of Coriolis forces for dipteran flies (house flies, mosquitoes, etc.). Instead of having four wings like other insects, dipterans have only two. Their hind wings have been reduced to tiny club-shaped halteres (see the figure) that beat at the same frequency as the fore wings. If their halteres are cut off, these flies become unstable in flight and soon crash to the ground. Pringle (*2*) explained how Coriolis forces on the halteres inform flies of rotations of their bodies, enabling them to fly stably. Sane *et al.* now find that hawk moths can do this with their antennae, although detection of aerodynamic as well as Coriolis forces may have a role.

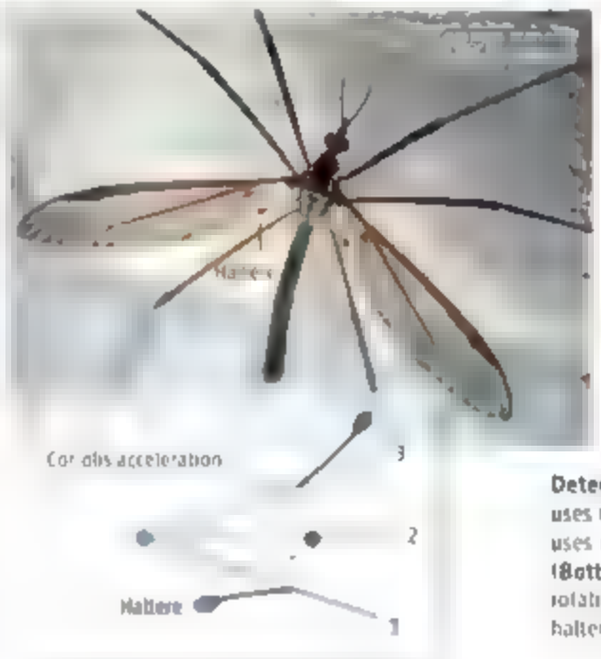
In the diagram, a rod representing a fly's body is rotating with constant angular velocity in the plane of the page. At the same time, a haltere hinged to the rod moves in a plane perpendicular to the paper. As the rod rotates from position 1 to position 3, the haltere swings from one side of it to the other, with constant angular velocity relative to the rod. Although both angular velocities are constant, the haltere has Coriolis acceleration, perpendicular to its plane of movement. Two effects contribute to this. First, the speed of the haltere is increasing as it moves further from the rod's axis of rotation. Second, the haltere's path curves to one side, as shown by the arrows. Coriolis forces must act on the haltere to give it these components of acceleration.

When a rotating-mass gyroscope is disturbed, the heavy wheel rotating at high speed strongly counteracts the disturbance. In contrast, the Coriolis forces on halteres are tiny, far too small to stabilize the fly directly. Instead, they cause slight bending of the stalk of the haltere, which is detected by sense organs, which in turn stimulate the fly to make the necessary adjustments to its wing beat.

Strepsiptera (tiny parasitic insects) have reduced fore wings that seem to work like halteres (*3*), but we did not know until now whether other insects use Coriolis forces to sense rotations. In their report, Sane and colleagues describe rotation detectors that work on the same principle in hawk moths. Like most other insects, moths have four wings and no halteres. They use their antennae in the way that flies use halteres. High-speed films of moths hovering in front of an artificial flower showed that the antennae vibrate at the wing beat frequency (about 27 Hz). When the flower was made to sway, the moth tracked its movements, and small deflections of the antennae, attributable to Coriolis forces, could be seen in the films. Separate electrophysiological experiments showed that sense organs at the bases of the antennae were sensitive enough to detect these deflections.

A simple experiment confirmed the importance of the antennae for stable flight. Intact moths hovered well in the flight chamber, but if their antennae were amputated near their bases, the moths were far more likely to crash to the ground or collide with the wall. When the antennae were glued back on, the moths' hovering ability was largely restored. Axons crossing the cuts had been severed, so the sensory information must have been provided by the organs at the bases of the antennae.

Animals need three-dimensional information about rotations. Vertebrates get this from three semicircular canals in each ear. One might expect each canal to be sensitive only to rotation in its own plane, but the situation is not quite so simple (*4*). Nevertheless, a set of three canals is sufficient to sense any



Detecting rotations. (**Top left**) A dipteran fly, which uses its halteres, and (**top right**) a hawk moth, which uses its antennae for rotation sensing during flight. (**Bottom left**) Diagram shows three positions during rotation of an insect and the Coriolis acceleration on the haltere or antenna.

PHOTO CREDITS: COMMON CRANE FLY, JANCO TAMS/WWW.FORESTRYIMAGES.ORG; PINE HAWK MOTH, STANISLAV KREJ/WWW.FORESTRYIMAGES.ORG

The author is in the Institute for Integrative and Comparative Biology, University of Leeds, Leeds LS2 9JT, UK. E-mail: r.m.alexander@leeds.ac.uk

rotation. One might not expect two halteres, or two antennae, to be capable of distinguishing three components of rotation, but they can, for a rather subtle reason.

The spectrum of Coriolis forces has peaks at the frequency of the haltere beat and at twice that frequency, and the relative magnitude of the peaks depends on the plane of rotation of the body. Thus, the sense organs in the stalk of the haltere experience a mix of two frequencies of mechanical stimulation, which changes as the plane of rotation changes, enabling each haltere to distinguish two components of rotation. The third possible component (in the plane of the haltere's beat) generates no out-of-plane force, but flies can still fly after one haltere has been amputated, suggesting that each haltere may be capable of distinguishing all three components. They could conceivably get the third one by sensing tension as well as bending in their stalks, but this would require the ability to measure tiny fluctuations in a much larger centripetal force (*5*). In any case, two halteres or antennae set at different angles, each capable of distinguishing two components of rotation, can together provide full information about rotations in three-dimensional space.

Now we would like to know whether other insects, as well as hawk moths, use their antennae to sense rotations. That capability seems potentially useful to any flying animal.

References

1. S. P. Sane, A. Dieudonné, M. A. Willis, T. L. Daniel, *Science* **315**, 863 (2006).
2. J. W. S. Pringle, *Philos. Trans. R. Soc. B* **233**, 347 (1948).
3. W. Pix, G. Nalbach, J. Zeil, *Naturwissenschaften* **80**, 371 (1993).
4. M. Muller, *Meth. J. Zool.* **50**, 279 (2000).
5. G. Nalbach, *J. Comp. Physiol. A* **173**, 293 (1993).

10.1126/science.1136840

ATMOSPHERIC SCIENCE

Pumping Up Surface Air

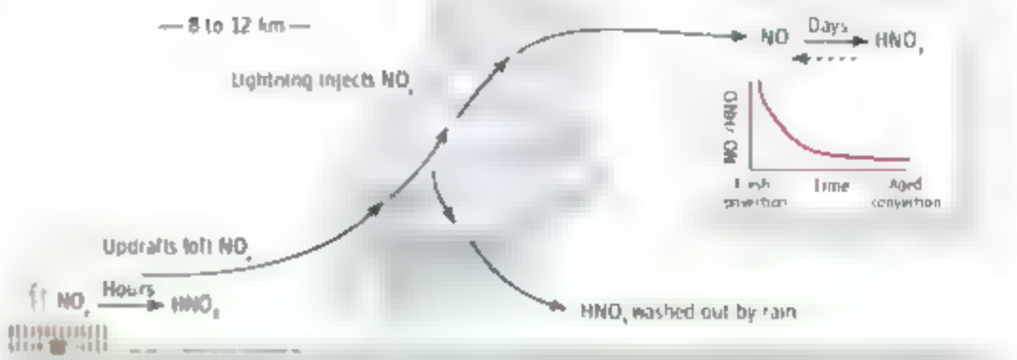
Towering cumulus clouds associated with thunderstorms enable the spread of pollution on intercontinental scales.

Lyatt Jaeglé

Many nitrogen, hydrocarbon, and sulfur pollutants involved in acid rain and photochemical smog are usually removed within hours or days near Earth's surface, but have much longer lifetimes if they reach the cold and dry upper troposphere at altitudes from 8 to 12 km. Towering cumulus clouds associated with thunderstorms (see the figure) can rapidly transport these pollutants to the upper troposphere. This process, combined with fast winds at high altitudes, allows the spread of pollution over intercontinental scales (*1*). Subsequent chemical transformation of these species in the upper troposphere leads to the production of ozone (*2*) and of aerosol particles (*3*), which affect global climate. In addition, convection moistens the upper troposphere, where water vapor has a large effect on climate and its response to increasing concentrations of greenhouse gases (*4*).

However, the rate at which the upper troposphere is flushed out and resupplied with fresh surface air cannot be measured directly. It is therefore difficult to evaluate the influence of deep convection on climate and air quality. On page 816 of this issue, Bertram *et al.* (*5*) use extensive observations obtained during a recent aircraft campaign to infer the turnover rate of the upper troposphere over the eastern United States and Canada during summer. These data will allow quantitative tests of models to assess how deep convection affects the composition of the upper troposphere.

The author is in the Department of Atmospheric Sciences, University of Washington, Seattle, WA 98195, USA. E-mail: jaegle@atmos.washington.edu



Tracking deep convection. A thunderstorm causes convective transport of NO_x from surface sources to the upper troposphere; lightning injects additional NO. Rain washes out the water-soluble HNO_3. High NO and low HNO_3 thus indicate fresh convection. As the cloud-processed air is transported downwind over the following days, NO_x is converted to HNO_3 at a rate determined by local chemistry. The NO_x/HNO_3 ratio thus provides a clock that allows the time since convection to be inferred.

Bertram *et al.* use measurements of nitrogen oxides (NO_x, which is the sum of NO and NO_2) and nitric acid (HNO_3) to quantify the influence of convection on the upper troposphere. The data were obtained by the NASA DC-8 research airplane, which sampled the atmosphere over eastern North America in July and August 2004 (*6*). Upper tropospheric air recently affected by upward motions in clouds usually contains high levels of NO_x but low levels of HNO_3, compared to the surrounding air (*7*). The NO_x originates from ground-level sources, such as electric utilities and cars, and from the lightning that accompanies convective clouds. NO_x is not very water-soluble and is thus efficiently transported through these clouds. In contrast, the very water-soluble HNO_3 is quickly removed by rain during convection. Once in the upper troposphere, NO_x is slowly oxidized to HNO_3 and HNO_3 regenerates NO_x until steady-state levels are reached 1 to 2 weeks later (see the figure).

Bertram *et al.* observed NO_x/HNO_3 ratios that were often 5 to 20 times as high as expected from steady-state equilibrium. They use this imbalance to calculate the time since convection occurred. Simultaneous measurements of hydrogen oxide radicals and other chemical parameters constrain the rate at which NO_x is oxidized to HNO_3. The authors use a photochemical box

model to predict the evolution of NO_x and HNO_3 after convection. This model allows them to relate the observed NO_x HNO_3 ratio to the time since convection. This "chemical clock" shows that about half of the air at altitudes between 7.5 and 11.5 km is less than 2 days old, and only 25% of the air has not been influenced by convection over the past 5 days.

The high time resolution of the measurements and the large region sampled by the research aircraft allow Bertram *et al.* to quantify the overall effects of convection. Their analysis implies that during summer, convection resupplies the upper troposphere with fresh surface air roughly every 5 to 10 days. This flood of surface air is injected by numerous individual thunderstorms. About 100,000 thunderstorms occur each year in the United States, mostly during summer.

What did this pervasive cloud pumping bring to the upper troposphere, in addition to copious amounts of NO_x? The authors found moderately enhanced levels of surface pollutants, such as sulfur dioxide and carbon monoxide, in freshly transported air. This fresh air also contained low concentrations of big particles, which had been scavenged during upward transport, and large concentrations of ultrafine particles, reflecting new particle formation. The chemical clock provides a means for timing the physical and chemical processing that occurs as the air ages, thereby providing new constraints on the rates of ozone production, particle formation, and mixing with surrounding air.

Because of limited computer power, global atmospheric models cannot simulate individual cloud systems. Instead, they use parameterizations of convection. The study by Bertram *et al.* provides a new benchmark for testing these representations of convection in a manner that does not depend on knowledge of surface emissions.

Future models may be able to represent the continental-scale effect of convection as captured in these new aircraft measurements. However, reproducing the actual location and intensity of convection on smaller scales remains a challenge that limits our ability to fully understand the extent of human influence on atmospheric composition.

References

1. E. E. Gidel, *J. Geophys. Res.* **88**, 6587 (1983).
2. K. E. Pickering *et al.*, *J. Geophys. Res.* **97**, 17985 (1992).
3. A. D. Clarke *et al.*, *J. Geophys. Res.* **103**, 16397 (1998).
4. D. Rind *et al.*, *Nature* **349**, 500 (1991).
5. T. H. Bertram *et al.*, *Science* **315**, 816 (2007); published online 4 January 2007 (10.1126/science.1134548).
6. H. B. Singh, W. H. Brune, J. H. Crawford, D. J. Jacob, P. B. Russell, *J. Geophys. Res.* **111**, D24S01 (2006).
7. R. R. Dickerson *et al.*, *Science* **235**, 460 (1987).

10.1126/science.1138988

MICROBIOLOGY

Bright Insight into Bacterial Gliding

A tank tread–like motion of successive adhesions provides the force needed to propel a rod-shaped bacterium.

Daniel B. Kearns

On page 853 of this issue, Mignot *et al.* provide evidence for a novel mode of bacterial locomotion (*1*). According to their model, a bacterium adheres to a solid surface at regular intervals along its cell body. Coordinated force is uniformly applied at focal adhesion complexes and the cell moves relative to these foci. Quantitative measurements and mathematical modeling predict that the zones of adhesion are organized on an internal helical track and that relative movement of the track could power a spiraling, corkscrew-like motion over a surface (see the figure).

The model is an explanation for the phenomenon of gliding motility, a form of bacterial movement that propels cells over a solid surface. While gliding, cellular motion is smooth and cells travel in straight lines, occasionally reversing in direction. Despite its observation in a wide variety of bacterial species, the mechanism of gliding motility has long remained a mystery, as gliding cells do not require external structures (such as flagella or pili) that could potentially serve as an obvious drive mechanism. With this in mind, Mignot *et al.* searched for motor proteins on the inside of *Myxococcus xanthus*, a slender, rod-shaped bacterium that moves by gliding. They focused on a coiled-coil protein called AglZ because mutant cells that lack the gene that encodes this protein are unable to glide (*2*). The bacteria were engineered such that yellow fluorescent protein was fused to AglZ, and when viewed under the microscope, functional AglZ appeared as brightly fluorescent intracellular foci. Through this approach, the researchers simply monitored the position of these foci as cells moved.

In watching the moving cells, Mignot *et al.* made several surprising observations. AglZ foci remain stationary with respect to the substrate, and thus a cell appears to move relative to the immobilized internal protein. This observation suggests that AglZ marks stationary foci (focal adhesion complexes) at which the cell anchors to a solid surface to push itself along. Consistent with this hypothesis, the cells do not form AglZ foci at positions where they do not contact the surface, such as during bending or rare flailing motions. Furthermore, when the lagging pole of a cell catches up to an AglZ cluster, the focus disperses and a new focus forms at the cell's leading pole. AglZ dispersal correlates with cellular relaxation, which suggests not only that these foci represent points of adhesion but also that force is applied at these points. The occurrence of AglZ foci at regularly fixed intervals implies that they associate with an intracellular helical track.

AglZ could participate in force generation. It is structurally similar to eukaryotic myosin, a motor protein that facilitates intracellular movement along cytoskeletal actin fibers. The periodicity observed for AglZ is consistent with the pitch of the bacterial cytoskeletal protein MreB, a homolog of eukaryotic actin, which forms a helix inside the cell (*3*). If complexes containing MreB and AglZ exist, they could theoretically generate force by a mechanism related to the myosin motors involved in eukaryotic cell division, white blood cell motility, and muscle contraction. Because both MreB and AglZ are found in the cytoplasm, the adhesion complex would necessarily have to traverse multiple layers of the cell envelope to interact with an external substrate. The focal adhesion–motor complex could be quite large, and numerous proteins have been implicated in gliding that could accommodate its assembly, structure, and function (*4*). Mysterious belt-like aggregates have been microscopically observed and biochemically purified from the outer layers of *M. xanthus* and might also participate in the gliding mechanism (*5*, *6*).

The author is in the Department of Biology, Indiana University, Bloomington, IN 47405, USA. E-mail: dbkearns@indiana.edu

To generate unidirectional movement, all components of the helical intracellular track must presumably travel in the same direction relative to the adhesion complexes and force must be applied uniformly over the helix. Track polarity must therefore also reverse when the cells reverse direction. To address this coordination, Mignot *et al* exploited the well-established fact that the frequency of direction reversals during gliding in *M. xanthus* is regulated by a series of signal transducers called the Frz proteins. They show that AglZ foci coalesce at the new leading pole of the cell concomitant with a change in direction, a correlation that is maintained even when the reversal frequency is radically altered by mutations in the Frz system. Therefore, it appears that the Frz signal transduction pathway may coordinate cellular reversals by controlling AglZ localization dynamics. Precisely how Frz controls the localization of AglZ and track polarity awaits explanation.

It is important to note that *M. xanthus* uses a second motor system—a retractile pilus—that is independent of the focal adhesion complexes but also powers movement over solid surfaces (see the figure). The Frz system also regulates this separate motor by controlling protein localization dynamics through a protein called FrzS (*7*). Thus, in addition to regulating the activity of two different motors, the Frz system must coordinate each such that they do not run simultaneously in opposite directions. How Frz integrates the two motors is unknown, but FrzS is remarkably similar in structure to AglZ, therefore, mechanistically distinct motors might be regulated by a common motif. Evidence suggests that FrzS is a motor regulator but is not required for pilus force generation (*7*). Similarly, it is possible that AglZ is not directly involved in force generation either, but rather is a regulator that colocalizes with the focal adhesion complexes.

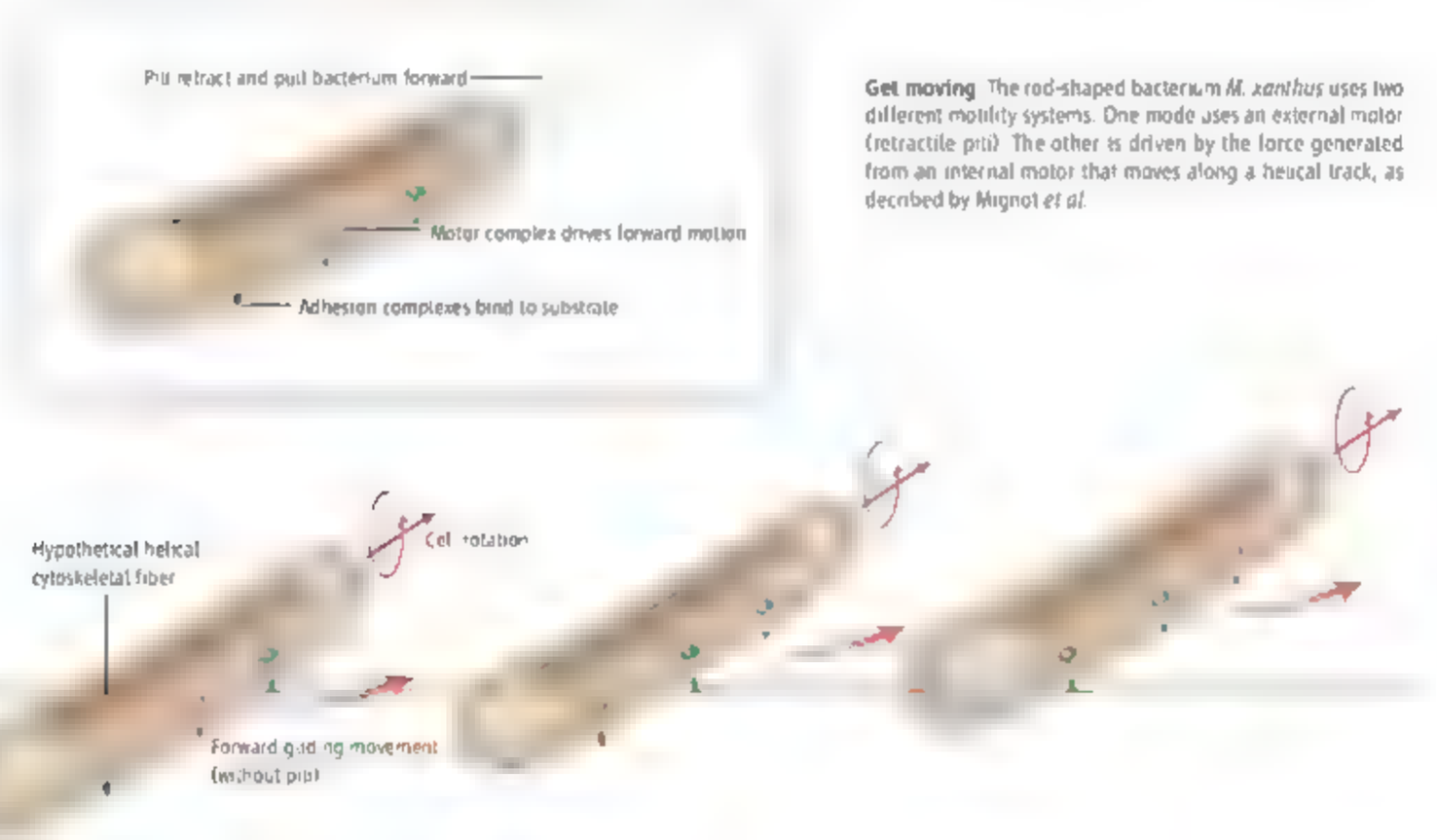
The model proposed by Mignot *et al* accommodates a number of observations regarding *M. xanthus* gliding motility. Specifically, focal adhesion complexes account for the lack of visible external drive structures during this type of movement and also explain the acute cell bending that occasionally occurs during gliding. Furthermore, the model is consistent with the previous observations of Sun *et al*, who demonstrated that gliding speed is independent of cell length and predicted that the gliding motor is distributed along the body of the cell and not localized to the cell pole (*8*). However, questions remain. What is the exact molecular nature of the intracellular helical track, and how is force generated? How is movement through the separate adhesion foci precisely coordinated? Finally, is AglZ regulatory or is it directly involved in force generation?

Gliding motility is found in many distantly related bacteria, and it seems likely that unique mechanisms for gliding have evolved independently (*9*). Nonetheless, Mignot *et al* provide a picture of one mechanism in particular, and the field of gliding motility now has a molecular handle by which to explore the larger phenomenon. Their observations may also provide insight into the origins of more complex eukaryotic motors.

References

1. T. Mignot *et al*., *Science* **315**, 853 (2007).
2. R. Yang *et al*., *J. Bacteriol.* **186**, 6168 (2004).
3. R. Carballido-López, *Microbiol. Mol. Biol. Rev.* **70**, 888 (2006).
4. P. Youderian, N. Burke, D. J. White, P. L. Hartzell, *Mol. Microbiol.* **49**, 555 (2003).
5. A. C. Burchard, R. P. Burchard, J. A. Kloetzel, *J. Bacteriol.* **132**, 666 (1977).
6. A. Friese, H. Reichenbach, H. Lünsdorf, *J. Bacteriol.* **179**, 1246 (1997).
7. T. Mignot, J. P. Merlie Jr., D. R. Zusman, *Science* **310**, 855 (2005).
8. H. Sun, Z. Yang, W. Shi, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 15178 (1999).
9. M. J. McBride, *Annu. Rev. Microbiol.* **55**, 49 (2001).

10.1126/science.1138995



Get moving The rod-shaped bacterium *M. xanthus* uses two different motility systems. One mode uses an external motor (retractile pili). The other is driven by the force generated from an internal motor that moves along a helical track, as decribed by Mignot *et al*.

CHEMISTRY

Can Droplets and Bubbles Think?

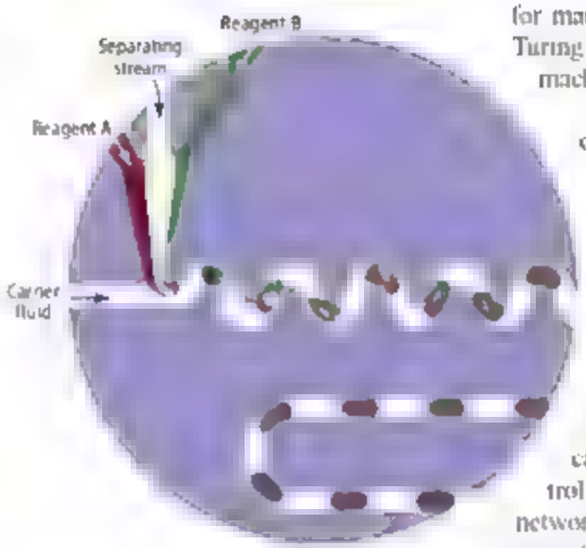
Irving R. Epstein

Bubbles flowing through narrow channels can be encoded with information and made to perform logic operations like those in a computer

How many of us, at the close of a well-spent evening, have stared into that last glass of beer or champagne and wondered what eternal truths lie in the rising bubbles before us? In this issue, Fuerstman *et al.* on page 828 (*1*) and Prakash and Gershenfeld on page 832 (*2*) report their use of microfluidic technology to construct streams of droplets (liquid-in-liquid) and bubbles (gas-in-liquid) that can encode and decode information or perform logical operations.

Since its emergence in the 1990s, microfluidics has become a powerful technique for a wide variety of applications in biotechnology, engineering, physics, and chemistry. By studying processes in channels with typical dimensions of tens to hundreds of micrometers, researchers can conduct controlled reactions while economizing on the consumption of possibly scarce materials. Flow in such channels can be characterized by two dimensionless numbers, the Reynolds number (*Re*), and the capillary number (*Ca*). When *Re* is low, inertial effects are negligible, and fluid flow is laminar and simple. In the narrow spaces of a typical microfluidic device, this is almost always the case.

The capillary number (*Ca*) characterizes the balance between surface tension and viscous forces. *Ca* determines the behavior of the bubbles or droplets when they reach a point at which a channel splits into two branches (*3*). For *Ca* less than a critical value that depends on the length of the bubble or droplet and the geometry of the channels, surface tension dominates, so the bubble holds together. Like the traveler in Robert Frost's poem *The Road Not Taken* (*4*), the bubble must choose a path. Although Frost's protagonist selects "the one less traveled by," a solitary bubble or droplet forced to settle on one of two paths will pick the channel that offers the lowest resistance to flow. Because the presence of a droplet or bubble increases the resistance of a channel, a second bubble arriving at a junction between two channels of similar but unequal length will take path B after the first takes the shorter path A. As the interval between droplets decreases, more complex sequences of path choices (e.g., AABAAB) can be generated. This simple principle provides the basis for the work reported by both of these groups.



Microfluidics plus chemistry. A microfluidic channel carries water droplets of volume 250 pl in a continuous stream of oil. [Adapted from (*18*)] The droplets act as microreactors in which the reagents are rapidly mixed (upper undulating region) and are then transported with no dispersion (lower smooth region). Such an arrangement can be used to control chemical reaction networks on a millisecond time scale.

Fuerstman *et al.* generate a series of droplets at periodically or aperiodically varying times and pass the resulting stream through a loop that offers the choice between two slightly different channels, thereby modifying the initial interdroplet time intervals and encoding an input signal. Remarkably, unlike Frost's traveler, who "doubted if I should ever come back," Fuerstman *et al.* are able to restore the initial series of intervals, thus decoding the signal, either by reversing the direction of flow or by passing the stream through a second loop.

Prakash and Gershenfeld take a more explicitly computational point of view, representing bits of information as bubbles in a channel, and construct circuits that not only carry out the functions of Boolean AND, OR, and NOT gates but also can be linked together to produce more elaborate arrangements capable of acting as counters, oscillators, or memory arrays. It is perhaps not totally farfetched to imagine a stream of bubbles and a set of rules for manipulating them as the equivalent of Turing's vision of a universal computing machine (*5*).

These demonstrations of the power of relatively simple two-component microfluidic flows make one wonder how far it is possible to go in constructing microfluidic "thinking devices." Can *in fluido* compete with *in silico* or *in cerebro*, at least for certain specialized applications? It is now possible to build microfluidic chips containing hundreds of addressable elements (*6*), and these elements can be made to perform important control and memory tasks (*7*). Microfluidic networks have already been used to solve mazes (*8*) and computationally difficult mathematics problems (*9*). Paradoxically, one obstacle to the practical use of microfluidic devices has been their lack of portability, but recent work (*10*, *11*) suggests ways in which this hurdle may be overcome.

Neurons and computer chips are certainly more complex than chemically inert droplets or bubbles. One way to enhance the capability of microfluidic devices is to combine them (see the figure) with the bistability, oscillations, and waves found in nonlinear chemical dynamics (*12*). On a macroscopic scale, the prototype Belousov-Zhabotinsky reaction has been used in devices that count (*13*), store images (*14*), solve mazes (*15*), and perform logical operations (*16*). Carried out in a microfluidic network, an appropriately chosen set of nonlinear reactions not only mimics but also provides insight into the mechanism of blood clotting (*17*). Although it is difficult to envision microfluidic systems in which the number of connections to a single element begins to approach the number of synapses to an individual neuron, adding the repertoire of chemistry to the flexibility and capability of rapid switching provided by microfluidics should make it possible to perform tasks of considerable complexity.

Will future James Bonds be equipped with microfluidic computers or encoders that fit into a shoe and are fueled by a bottle of spring water (or premium vodka)? Probably not, but the possibility of carrying out an impressive variety of tasks usually associated with brains or computers will undoubtedly inspire many more creative uses of microfluidics, both in

The author is in the Department of Chemistry and the Volen Center for Complex Systems, Brandeis University, Waltham, MA 02454–9110, USA. E-mail: epstein@brandeis.edu

CREDIT: R. KRAGHOLM/UNIVERSITY OF CHICAGO

systems like the ones explored in these two reports and in more complex arrangements that exploit chemistry as well

References

1. M. J. Fuerstman, P. Garstecki, G. M. Whitesides, *Science* **315**, 828 (2007).
2. M. Prakash, N. Gershenfeld, *Science* **315**, 832 (2007).
3. O. R. Link, S. L. Anna, D. A. Weitz, H. A. Stone, *Phys. Rev. Lett.* **92**, 054503 (2004).
4. R. Frost, *Mountain Interval* (Henry Holt, New York, 1921), pp. 9 and 10.
5. A. M. Turing, *Proc. London Math. Soc.* **42**, 230 (1936).
6. T. Thorsen, S. J. Maerkl, S. R. Quake, *Science* **298**, 580 (2002).
7. A. Grossman, M. Enzelberger, S. R. Quake, *Science* **300**, 955 (2003).
8. M. J. Fuerstman *et al.*, *Langmuir* **19**, 4714 (2003).
9. D. T. Chiu, E. Pezzoli, H. K. Wu, A. D. Stroock, G. M. Whitesides, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2963 (2001).
10. P. Garstecki, M. J. Fuerstman, M. A. Fischbach, S. K. Sia, G. M. Whitesides, *Lab Chip* **6**, 207 (2006).
11. J. Moorthy *et al.*, *Electrophoresis* **25**, 1705 (2004).
12. I. R. Epstein, J. A. Pojman, *Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos* (Oxford Univ. Press, New York, 1998).
13. J. Goredki, K. Yoshikawa, Y. Igarashi, *J. Phys. Chem. A* **107**, 1664 (2003).
14. A. Kamimaga, V. K. Vanag, I. R. Epstein, *Angew. Chem. Int. Ed.* **45**, 3087 (2006).
15. O. Steinbock, A. Toth, K. Showalter, *Science* **267**, 868 (1995).
16. O. Steinbock, P. Kettunen, K. Showalter, *J. Phys. Chem.* **100**, 18970 (1996).
17. M. K. Rumyon, B. L. Johnson-Kerner, R. F. Ismagilov, *Angew. Chem. Int. Ed.* **43**, 1531 (2004).
18. H. Song, J. D. Tice, R. F. Ismagilov, *Angew. Chem. Int. Ed.* **42**, 768 (2003).

10.1126/science.1138325

CELL BIOLOGY

Cellular Demolition and the Rules of Engagement

Does cell death occur by pulling the plug on Bcl-2 life support or pulling the trigger on Bax activation?

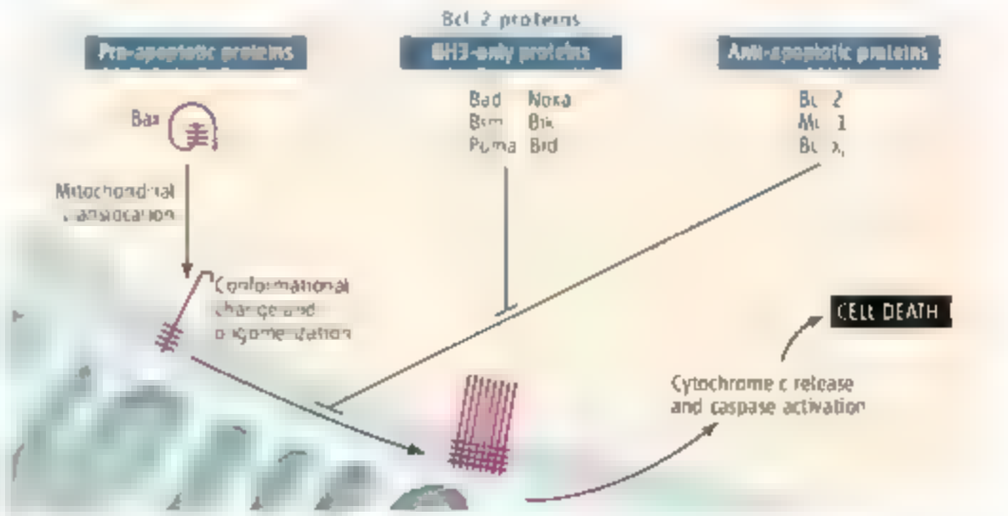
Richard J. Youle

Execution of cellular suicide, or apoptosis, is dramatically under the control of one group of proteins, the Bcl-2 family. But within this family, these factions are at odds with each other as they rival to promote or block cell death. Anti-apoptotic members act to restrain pro-apoptotic members. More of an enigma has been the activity of "BH3-only proteins," family members that share a single motif (a BH3 domain) with other Bcl-2 proteins. BH3-only proteins can bind to different subsets of the Bcl-2 family, but whether they assist cell suicide by activating pro-apoptotic members or inhibiting anti-apoptotic family members has been debated. On page 856 of this issue, Willis and colleagues (*1*) conclude that BH3-only proteins work solely by thwarting anti-apoptosis Bcl-2 proteins, thus settling the controversy. We now understand one way to start the cell's death machinery.

Bcl-2 regulates tissue health by inhibiting apoptosis. Bax (and its homolog Bak) has the opposite effect of promoting apoptosis despite a surprisingly similar three-dimensional structure. It is now accepted that anti-apoptotic family members, including Bcl-2, Bcl-x_L, and Mcl-1, restrain Bax and Bak activity. Once activated, Bax and Bak induce permeabilization of the mitochondrial membrane, thereby allowing cytochrome *c* release and the activation of caspases (which catalyze protein degradation). This cascade of events ultimately leads to breakdown of the cell (see the figure).

Bax was originally identified as a Bcl-2 binding protein and was thought to be inhibited by this interaction (*2*). However, contrary to this "rheostat model," in healthy cells, Bax is not bound to Bcl-2 but exists as a monomer in the cytosol (*3*). How can anti-apoptotic Bcl-2 family members keep pro-apoptotic Bax in check without direct interaction? One solution may center on the actions of BH3-only proteins.

Cell-free models of Bax activation, which do not fully reflect the in vivo situation, show that pro-apoptotic Bax acts synergistically with peptides corresponding to the BH3 domain of only a few BH3-only proteins, notably those from Bid and Bim, to induce cytochrome *c* release from isolated mitochondria (*4*–*6*). Bid and Bim are thus referred to as "activator proteins." This scenario is consistent with the simple model in which BH3 peptides bind to Bax and cause a conformational change and oligomerization of Bax, thereby activating the mitochondrial cell death pathway. How then do the other BH3-only proteins that do not bind to Bax (so-called "nonactivating" proteins) induce apoptosis? One possibility is that activator proteins could be sequestered by Bcl-2. The activators could be freed from this association if displaced by nonactivating BH3-only proteins. The released activators could then bind to Bax and trigger cell death (*5*). One problem of this model is that activator proteins (Bim and Bid) do not detectably bind to Bax. However, recent work shows that stabilizing the BH3 domain structure (α helices) of Bid and Bim allows the peptides to bind to Bax, keeping the theory plausible (*7*).



Restraining orders. Anti-apoptotic Bcl-2 family members block the translocation of pro-apoptotic Bax perhaps by directly binding and inhibiting Bax on mitochondria following a conformational change. BH3-only proteins promote apoptosis by binding to and inhibiting anti-apoptotic Bcl-2 family protein activity.

The author is in the Surgical Neurology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA. E-mail: youler@ninds.nih.gov

In contrast to the cell-free models, Willis *et al.* show that Bid and Bim are not at all required for apoptosis in vivo. Using mice lacking Bid and Bim, the authors show that nonactivating proteins still induce apoptosis in vivo. They also confirm that Bim and Bid fail to bind pro-apoptotic Bak and Bax. Thus, BH3-only proteins appear able to promote apoptosis solely by inhibiting Bcl-2 and anti-apoptotic homologs to promote apoptosis, without displacing activator proteins.

What then activates Bax and Bak to set the mitochondrial cell death pathway in motion? The situation for Bak could be more straightforward than for Bax. In contrast to cytosolic and monomeric Bax, Bak is normally bound to mitochondria and constitutively bound to Mcl-1 and Bcl-x_L, anti-apoptotic Bcl-2 family members (*8*, *9*). Thus, simple relief from inhibition through binding of BH3-only proteins could unleash Bak to oligomerize and induce apoptosis as originally proposed in the rheostat model. However, mutant forms of Bak less able to bind Bcl-x_L and Mcl-1 still require an unknown activation step to induce apoptosis (*6*), indicating that simply removing inhibition by Mcl-1 and Bcl-x_L is not enough.

How cell death–promoting Bax is activated remains unknown. It translocates from the cytosol to the mitochondria during apoptosis, and overexpression of Bcl-2 somehow prevents this relocation. It may be that cellular stress alters the conformation of Bax to a semi-activated state, some of the Bax then moves to mitochondria, where anti-apoptotic Bcl-2 family members bind and inhibit it (see the figure). Similarly, nascent Bak oligomers on mitochondria are bound by Bcl-2 and thereby blocked from further oligomerization and from inducing apoptosis (*10*). These semiactivated forms of Bax and Bak could be relieved of Bcl-2 inhibition by BH3-only proteins to allow completion of translocation, oligomerization, and cytochrome *c* release.

Complicating the issue further are experiments indicating that Bcl-2 must change conformation on the mitochondria to inhibit Bax (*11*). Thus, activated forms of Bcl-2 may block activated forms of Bax, again consistent with the rheostat model except that inhibition occurs downstream of the elusive Bax and Bcl-2 activation steps. Two central helices of Bax and Bcl-2 become embedded deeply in the mitochondrial membrane during activation, thereby unfolding and undoubtedly destroying domains that harbor binding pockets for BH3 domains. Elucidating how BH3-only proteins relieve inhibition of membrane-inserted Bax by Bcl-2 requires fresh insight into the structures of the membrane-embedded forms.

Willis and colleagues make clear that BH3-only proteins can promote apoptosis exclusively by inhibiting anti-apoptotic Bcl-2 family members (see the Table). Moreover, particular BH3-only proteins engage particular anti-apoptotic Bcl-2 family members, allowing tissue-specific regulation and stress-specific cellular responses. Beyond the issue of Bax and Bak activation remains the central question of how they stimulate mitochondrial membrane permeabilization, a process that has defied intense research efforts so far.

References

1. S. N. Willis *et al.*, *Science* **315**, 856 (2007)
2. S. J. Korsmeyer, J. R. Shutter, D. J. Veis, D. E. Merry, Z. N. Oltvai, *Semin. Cancer Biol.* **4**, 327 (1993)
3. Y. T. Hsu, R. J. Youle *J. Biol. Chem.* **273**, 10777 (1998)
4. T. Kuwana *et al. Mol. Cell* **17**, 525 (2005)
5. M. Certo *et al. Cancer Cell* **9**, 351 (2006)
6. H. Kim *et al., Nat. Cell Biol.* **8**, 1348 (2006)
7. L. D. Walensky *et al. Mol. Cell* **24**, 199 (2006)
8. A. Cuconati, C. Mukherjee, D. Perez, E. White *Genes Dev.* **17**, 2922 (2003)
9. S. N. Willis *et al., Genes Dev.* **19**, 1294 (2005)
10. S. C. Ruffolo, G. C. Shore, *J. Biol. Chem.* **278**, 25039 (2003)
11. R. J. Dlugosz *et al. EMBO J.* **25**, 2287 (2006)

10.1126/science.1138870

OCEANS

Picoplankton Do Some Heavy Lifting

Tiny unicellular plants called picoplankton play an unexpectedly large role in the transport of organic matter to the deep ocean

Richard T. Barber

Over the past 25 years, understanding of the food web structure of open-ocean ecosystems has changed dramatically. It is now clear that the traditional food web, which is a chain from diatoms through crustaceans to fish, is limited to high-biomass diatom blooms that occur only intermittently in the open ocean. A background food web based on very small unicellular plants (picoplankton, with diameters <5 μm) and small protozoan animals is present in all regions of the ocean more or less continually throughout the year (*1*–*5*). Biological oceanographers have assumed that this "microbial food web" plays a minor role in important oceanic processes such as fish production and sequestration of carbon in the deep ocean.

On page 838 of this issue, Richardson and Jackson (*6*) challenge this assumption by showing that autotrophic (photosynthetic) picoplankton are a source of organic carbon for large zooplankton such as copepods and for the particulate organic carbon pool that fuels the flux of particles sinking to the deep ocean (see the figure).

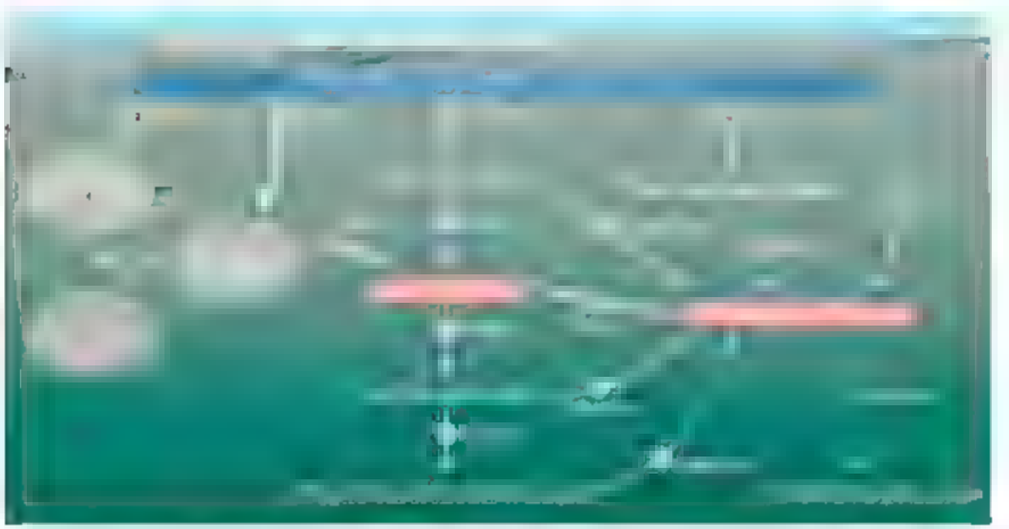
The authors used inverse and network analyses to establish that picoplankton transfer much more newly synthesized carbon through the food web than had previously been assumed. This finding differs from the present dogma of biological oceanography, which is that the transfer of picoplankton carbon through the food web is disproportionately low relative to their net primary productivity (NPP) (*7*), whereas the contribution of diatoms is disproportionately high (*8*–*11*). Richardson and Jackson do not dispute the importance of diatom blooms, when they occur, but they show convincingly that picoplankton can also be an important source of new organic carbon for large zooplankton and detritus production.

The productivity observations that were the starting point for Richardson and Jackson's analyses came from the U.S. Joint Global Ocean Flux Study (JGOFS) in the equatorial Pacific in 1992 and in the northern Arabian Sea in 1995. These cruises sampled two distinct El Niño Southern Oscillation (ENSO)

The author is in the Nicholas School of the Environment and Earth Sciences, Duke University, Beaufort, NC 28516, USA. E-mail: rbarber@duke.edu

states in the Pacific and three monsoon conditions in the Arabian Sea. Picoplankton NPP was a substantial fraction (between 60 and 90%) of total NPP at all sites except those sampled during the Southwest Monsoon. Even in the nutrient-rich upwelling of the Southwest Monsoon (*12*), picoplankton NPP was 30 to 50% of total NPP. The new results reported by Richardson and Jackson are thus relevant not only for the well-stratified, low-nutrient conditions where picoplankton are usually dominant (*13*, *14*), but also for high-nutrient conditions.

The thesis advanced by Richardson and Jackson challenges the mainstream concept of the role played by picoplankton export, but there may be less disagreement than is appar-



The picoplankton food web. This oceanic food web based on picoplankton shows the paths of organic carbon flux determined by Richardson and Jackson (*6*). On the left is the classical "microbial loop" (gray). The two red boxes (large zooplankton and particulate organic detritus) are two carbon pools that, according to Richardson and Jackson, receive substantial export of picoplankton carbon. This new information suggests that the role of picoplankton in carbon export and fish production needs further investigation in both observations and models.

ent at first reading. Landry (*14*) has shown that unicellular animals that graze on picoplankton have high potential growth rates and can keep pace with picoplankton growth, thereby efficiently consuming the extra picoplankton biomass that is produced during the onset of favorable growth conditions. As a result of this balance between growth and grazing, picoplankton do not accumulate in a high-biomass bloom. Without such blooms, the amount of picoplankton biomass exported to the deep sea is assumed to be minimal. Evidence for Landry's balance hypothesis (*14*) comes from an ingenious dilution-incubation experiment (*15*) that determines both the growth rate of the autotrophs and their mortality rate due to grazing.

However, the dilution-incubation experiment is not designed to determine all possible paths of picoplankton carbon, such as those shown in paths 3 to 5 in the figure. For example, the carbon flux due to the consumption of organic aggregates by large zooplankton (path 7) cannot be quantified at present, nor have there been many studies quantifying the impact of tunicate and pteropod grazing on picoplankton (path 4). Without organic aggregates and tunicate grazers in the short-term dilution-incubation experiments, the production of particulate organic carbon detritus by aggregate disintegration or tunicate defecation cannot be determined.

In contrast, the inverse and network analyses used by Richardson and Jackson were designed to determine all possible carbon paths and their variability as a function of monsoon and ENSO state at a variety of locations. Over the time scale resolved by their analyses (weeks to months), the picoplankton produced more organic carbon than was eaten by their grazers, a portion of this new picoplankton carbon was exported to the large zooplankton and detritus pools (*6*). Richardson and Jackson suggest that the export fluxes they determined probably involved the formation of organic aggregates from picoplankton (path 3), consumption of those aggregates by large zooplankton (path 7), and the production of fecal material by pelagic tunicates grazing on the picoplankton (paths 4 and 8).

If there is indeed substantial export of picoplankton carbon to large zooplankton and detritus, then some concepts about oceanic biogeochemical cycling will need to be revised. The abundance of picoplankton tends to be relatively low, but they are continuously present over the vast areas of tropical and subtropical oceans and, in a time-space summation, picoplankton NPP adds up to a major source of marine NPP (*1*–*5*). Furthermore, under favorable growth conditions, both picoplankton and diatoms increase in absolute biomass and productivity (*11*). The magnitude of global NPP by picoplankton is probably even higher than previously claimed, and the new interpretation of Richardson and Jackson will thus be of great interest to ocean ecologists, biogeochemists, and modelers investigating how climate change will affect carbon cycling or oceanic fish production.

As Richardson and Jackson point out, most ocean ecosystem models do not give much or any weight to picoplankton export as a source of sinking particulate organic carbon. On the other hand, the models cited by Richardson and Jackson do include the link between picoplankton, protistan grazers, and large zooplankton (see paths 2 and 6 in the figure) that this analysis shows to be important. It should therefore be possible to use these models for examining the consequences of the new flux rates determined by Richardson and Jackson.

Coupled ocean-atmosphere models forced with rising CO_2 indicate that upper-ocean stratification and stability will increase and mixing will decrease in the next 50 years (*16*). These conditions increase the dominance of picoplankton (*13*–*15*) relative to diatoms and other microplankton. It is reasonable to assume that NPP by picoplankton will increase and NPP by larger microplankton will decrease, especially at low and mid-latitudes. The issues raised by Richardson and Jackson thus merit careful scrutiny and confirmation by other researchers.

References and Notes

1. J. B. Waterbury *et al.*, *Nature* **277**, 293 (1979).
2. W. K. W. Li *et al.*, *Science* **219**, 292 (1983).
3. F. Azam *et al.*, *Mar. Ecol. Prog. Ser.* **10**, 257 (1983).
4. T. Platt *et al.*, *Nature* **300**, 701 (1983).
5. S. W. Chisholm *et al.*, *Nature* **334**, 340 (1988).
6. T. L. Richardson, G. A. Jackson, *Science* **315**, 838 (2007).
7. Net primary productivity is defined as the total newly synthesized organic carbon minus the organic carbon used by the producing plant for their own respiration.
8. J. Ryther, *Science* **166**, 72 (1969).
9. D. Cushing, *J. Plankton Res.* **11**, 1 (1989).
10. R. Iverson, *Limnol. Oceanogr.* **35**, 1593 (1990).
11. R. T. Barber, M. R. Hiscock, *Global Biogeochem. Cycles* **20**, GB4503 (2006).
12. R. T. Barber *et al.*, *Deep-Sea Res. II* **48**, 1127 (2001).
13. M. R. Landry *et al.*, *Deep-Sea Res. II* **45**, 2353 (1998).
14. M. R. Landry *et al.*, *Limnol. Oceanogr.* **42**, 405 (1997).
15. M. R. Landry, R. P. Hassett, *Mar. Biol.* **67**, 283 (1982).
16. J. L. Sarmiento *et al.*, *Global Biogeochem. Cycles* **18**, GB3003 (2004).

17. I thank the NSF for its generous support of the U.S. JGOFS.

10.1126/science.1137438

page 790

CREDIT: DAVID PAUL MORRIS/GETTY

INTRODUCTION

Energy for the Long Haul

PERHAPS THE GREATEST CHALLENGE IN REALIZING A SUSTAINABLE FUTURE IS energy consumption. It is ultimately the basis for a large part of the global economy, and more of it will be required to raise living standards in the developing world. Today, we are mostly dependent on nonrenewable fossil fuels that have been and will continue to be a major cause of pollution and climate change. Because of these problems, and our dwindling supply of petroleum, finding sustainable alternatives is becoming increasingly urgent. This special issue focuses on some of the challenges and efforts needed to harness renewable energy more effectively at a sufficient scale to make a difference and some of the people who are working on these problems. As introduced in the first News article (p. 782), the Editorial by Holdren (p. 737), and the Perspective by Whitesides and Crabtree (p. 796), many of the outstanding questions require major research efforts in underfunded areas.

Much of the focus on sustainable energy is aimed at different ways of tapping into the most abundant renewable resource: solar energy. Lewis (p. 798) points out that the direct conversion of sunlight with solar cells, either into electricity or hydrogen, faces cost hurdles independent of their intrinsic efficiency. Ways must be found to lower production costs and design better conversion and storage systems. In the short term, utilization of biomass relies mainly on sugar fermentation. Goldemberg (p. 808) discusses how Brazil's use of ethanol from sugarcane has greatly reduced its need for imported oil. Many long-term goals have been set for biomass utilization. For example, the European Union (EU) hopes to produce a quarter of its transportation fuels from biomass by 2030, as discussed by Himmel *et al.* (p. 804). Better ways are also needed for processing the available sugars, and conversion to higher alcohols or even alkanes is desirable. Stephanopoulos (p. 801) explores the options afforded by reengineering biosynthetic pathways in microbes.

How we tackle energy problems will turn on a number of policy issues. Potočnik (p. 810) discusses how the EU is setting targets and allocating funding for alternative energy. Finally, Schrag (p. 812) explores the feasibility of sequestering carbon dioxide from fossil fuel use and our technological readiness and willingness to implement such schemes.

The News section profiles national lab directors, computer modelers, captains of industry, and bench scientists who are writing the early chapters of the next book on energy research. Some of them are developing better plants to grow as fuel or ways to convert them into ethanol. Others are developing catalysts to extract hydrogen from water or generate electricity from hydrogen. What they all share is a desire to find new ways to power the future. ScienceCareers.org takes a look at three young private-sector scientists who are on their first steps to careers in energy R&D: a consultant helping Israel meet its obligations under the Kyoto Protocol, a former particle physicist designing solar energy systems, and a Ph.D.-level engineer integrating sustainable electricity supplies into the power grid.

– PHIL SZUROMI, BARBARA JASNY, DANIEL CLERY, JAMES AUSTIN, BROOKS HANSON

Sustainability and Energy

CONTENTS

See also related Editorial page 737; Science Careers section page 868; Podcast; www.sciencemag.org/sciext/sustainability

Science

NEWS

A Sustainable Future, If We Pay Up Front

The oil shocks of the 1970s sparked a short-lived golden age of energy R&D. Today's energy concerns haven't caused a similar bonanza, yet

IT'S BEEN A YEAR SINCE PRESIDENT GEORGE W. Bush called on researchers to help the United States shake off its "addiction" to oil and its reliance on foreign sources of energy. "By applying the talent and technology of America, this country can dramatically improve our environment, move beyond a petroleum-based economy, and make our dependence on Middle Eastern oil a thing of the past," he said. Since then, talk of breaking the petroleum habit has intensified. Political instability in key regions that supply oil and natural gas and oil prices near an all-time high have made energy the issue of the moment, with newspapers, magazines, and TV news programs extolling the virtues of rooftop solar cells, wind turbines, biofuels, and hybrid-energy cars.

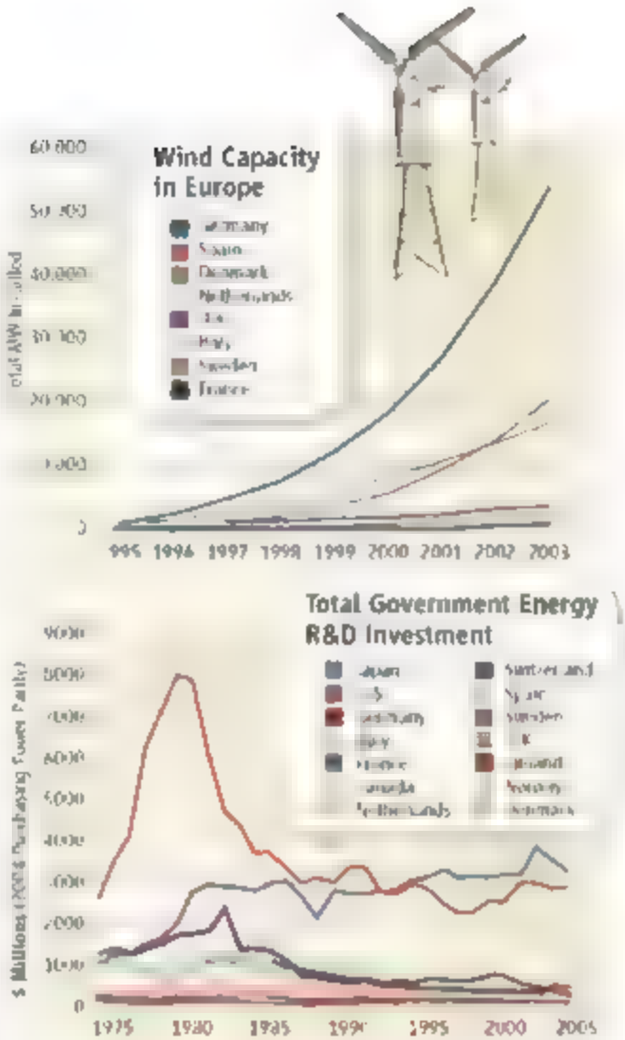
All this is reminiscent of the late 1970s, when oil shocks prompted a surge of interest in alternatives to conventional sources of oil. But there are some significant differences. Today's concern is driven by worries about climate change in addition to fears of energy insecurity. And whereas the 1970s oil shocks prompted a gusher of new funding for energy-related R&D, this time around, research dollars—or euros—have yet to begin flowing in earnest. The science has also changed over the past quarter-century. Now geneticists and synthetic chemists are tweaking agricultural crops to make better liquid fuels, materials scientists are wringing more efficiency out of photovoltaic cells, and computer modellers are redesigning nuclear reactors from the ground up. In the following news pages, we profile some of the people who are following these new paths in energy research.

Energy-policy experts and several recent studies agree that if we are to make any substantial change to our energy supply, huge increases in funding will be needed. But prospects for sharp increases in energy research seem unlikely. "There certainly should be [more funding]," says carbon sequestration expert Robert Socolow of Princeton University. Although there are some signs of new money in targeted areas of alternative-energy research both in the United States and elsewhere, "I don't see major increases because of the federal deficit," says physicist Ernest Moniz, an energy-policy specialist at the Massachusetts Institute of Technology in Cambridge and former undersecretary at the U.S. Department of Energy (DOE). "There are tough fiscal times ahead." Indeed, in the United States, the anticipated yearlong budget freeze could delay even some of the projects that have been tapped for increases (*Science*, 15 December 2006, p. 1666).

These dim funding prospects follow a period of relative decline in energy R&D. Paul Runci of the Pacific Northwest National Laboratory in Richland, Washington, produced an analysis last year of energy R&D investment in 11 nations that together provide 95% of the energy research funding in industrialized countries. Almost all of those countries steadily increased funding until the late 1970s to mid-1980s, when levels tapered off until the mid-1990s. They have remained largely stagnant ever since (see figure, p. 783). The biggest falls have been in funding of nuclear and fossil fuel R&D (see figure, p. 783). In some countries, fossil fuel research is less than 10% of what it was at the peak. The exception is Japan, which kept increasing funding until this decade. Runci blames this drop on a decade of low oil prices and oversupply, changing perceptions of nuclear power following the accidents at Three Mile Island in Pennsylvania and at Chornobyl, Ukraine, and public policies that have shifted R&D responsibility from the public to the private sector.

Industry has so far failed to take up the slack, while government money has dwindled. The U.S. energy industry is a factor of 10 below the average R&D intensity of American business. The shortage of both public and private funding is having a serious impact on new blood coming into this area of research. "Fields saturate [with older researchers]," says Socolow. "Energy keeps not getting enough new recruits."

CREDIT: GETTY IMAGES



SOURCES (TOP TO BOTTOM): EUROPEAN UNION; PEW; GALLAGHER ET AL. ANN. REV. ENVIRON. RESOUR. 31, 193 (2006)

There are some signs of a revival, however. Runci's report points to energy-conservation research, which grew strongly in the United States and Japan during the late 1990s and into this decade. Runci also singles out the growth of wind energy in Europe. Concern about carbon emissions and a desire to become world leaders in renewables technology spurred the European Union (E.U.) and some national governments to invest in wind energy R&D as well as provide subsidies and tax incentives to energy companies to set up wind farms. The result has been a phenomenal growth in installed wind capacity (see figure, top right) that is now equivalent to 50 coal-fired power stations. The E.U.'s renewables industry now has an annual turnover of $20 billion, half the world market, and employs 300,000 people.

In his speech last year, Bush voiced new enthusiasm for novel forms of energy technology and announced the launch of his Advanced Energy Initiative. He subsequently proposed a 22% increase in DOE funding for clean energy technology in 2007, focusing on improving the efficiency of cars as well as expanding the use of biofuels and developing fuel cells, and on electricity-generating technologies including clean coal, advanced nuclear power, solar, and wind. These areas received $1.7 billion in funding during 2006, but some may get caught in the mid-pace budget freeze.

One of the most eagerly awaited DOE initiatives is a plan to build two new bioenergy research centers to carry out basic research on microbial and plant systems with the aim of converting inedible plant fibers such as cellulose into ethanol for biofuels. With $125 million over 5 years on the table for each center, universities, national labs, and other institutions are vying to host them. A decision is expected in the summer. The oil company BP has launched a similar scheme, offering $500 million over 10 years to an institution to run an Energy Biosciences Institute (see p. 790). A decision on the site, either in the United States or the United Kingdom, is expected by the time this issue is published. "We've come to realize that the intersection of biosciences and energy is a very powerful one," says Steven Koonin, BP's chief scientist.

The European Union is also backing new energy technologies. In the latest 7-year Framework research program—approved late last year—$1.5 billion will be spent on renewables and energy efficiency, an increase of 40% over the previous Framework program and more than 50% of all non-nuclear energy funding. In the United Kingdom, meanwhile, a government energy review completed last summer promised a National Institute for Energy Technologies, with a budget in the region of $200 million per year.

All of these initiatives are giving energy research a much-needed shot in the arm, but a number of influential reports are calling for increases in energy R&D on an entirely different scale. The G8 meeting of world leaders at Gleneagles, U.K., in July 2005 called on the International Energy Agency (IEA) to map out scenarios for a future clean energy supply. According to the IEA report, published last year, "there is an acute need to stabilize declining budgets for energy-related R&D and then increase them." Nicholas Stern, a chief economic adviser to the U.K. government, was more specific. In his influential review on the economics of climate change published last October, he forecast that the impact of doing nothing to counter global warming could amount to between 5% and 20% of global gross domestic product, whereas efforts costing 1% of GDP could counter the worst effects. Those efforts would include at least doubling the current amount of energy R&D funding and a fivefold increase in support for the deployment of new technologies.

An ambitious report released in 2005 by the U.S. National Academies entitled *Rising Above the Gathering Storm* aimed to identify actions to help the United States compete and prosper in the 21st century. Among a number of recommendations, it advocated the creation of a new agency within DOE, akin to the Defense Advanced Research Projects Agency, to fund energy R&D. This new body, dubbed ARPA-E, should start with $300 million a year and build up to $1 billion annually over 5 or 6 years. A bill to set up ARPA-E was presented to Congress in December 2005 by Representative Bart Gordon (D-TN). Gordon, now the House science committee chair, reintroduced the bill last month. Moniz says that research he and his colleagues have conducted for reports on the future of nuclear energy and coal also suggests that a doubling of research spending is necessary.

Ultimately, it will be politics at the highest level that will dictate the fate of energy research. If oil supplies are further threatened by geopolitics, or new evidence of human-induced climate change is sufficient to convince the skeptics, energy research may again prosper. But will it continue to do so if the price of oil later drops again? "That remains to be seen," says Koonin.

–DANIEL CLERY

PROFILE: STEVEN CHU

Steering a National Lab Into the Light

Steven Chu is on a crusade to make solar power work. His weapon is Lawrence Berkeley National Laboratory

Drive past the guardhouse at the entrance to Lawrence Berkeley National Laboratory (LBNL), and you travel about 250 meters on Chu Road before it splits off onto other tentacles of the lab overlooking San Francisco Bay. Despite its short length, it's a fitting entree. The road is named for 1997 physics Nobelist Steven Chu, who became the lab's sixth director 2 years ago and ever since has been working to steer its research in a new direction. His focus, a sweeping agenda aimed at making solar energy a practical, commercial, and world-changing reality.

Since taking his new job, Chu, 58, has traveled the globe making the case for solar power. Some of the refrains are familiar: retreating glaciers, more-damaging storms, and sea-level rise. Some, though, are less well known, such as how increased water evaporation from soils could jeopardize farming in the Midwestern United States. "Sustainable carbon-neutral energy is the most important scientific challenge we face today," Chu is fond of saying.

On the supply side, Chu says current carbon-neutral energy sources face a host of problems. Wind, nuclear power, and biofuels are unlikely anytime soon to meet all of the needs of a civilization that in 2005 was using energy in all forms at a rate of about 16 trillion watts, with roughly 80% of it coming from fossil fuels. Solar power also has limitations, but the sun beams more energy toward Earth in an hour than all 6 billion of us use in a year. The challenge is finding ways to capture and store that energy that are cheap and efficient.

To meet the challenge, Chu has launched three separate initiatives to yoke at least part of LBNL research to a solar agenda. The first is a proposal to team up with fellow national labs Sandia and Lawrence Livermore and build one of two $125 million Bio Energy Institutes to be funded by the Department of Energy (DOE) that, over 5 years, will propel advances in nanotechnology and synthetic biology into better ways of making biofuels. Second is to win a $500 million pot put up by the oil giant BP for a biofuels institute (see p. 790). Finally, an initiative called Helios will look beyond bioenergy to include topics such as nanotech-based photovoltaics and fuel-generating catalysts. Although none of the proposals had been funded by the time *Science* went to press, Chu already has verbal commitments from private sources of about $50 million and up to $70 million from the state of California. "You try to make some run and get some funding," Chu says.

If that funding comes, Chu expects little trouble getting scientists to follow and little opposition within LBNL, because the programs would expand the lab's current $325 million annual budget rather than cut into existing programs. The new programs also wouldn't affect the lab's four national user facilities—a synchrotron, an electron microscopy center, a nanofabrication facility, and a supercomputing center—other than to encourage users to target their research on improving solar power generation. So far, Chu's ambitions are getting a warm reception both inside and outside the lab. "I think it's great," says physicist Nate Lewis of the California Institute of Technology in Pasadena, adding: "If you need to lay more golden eggs, the only proven way is to buy more chickens." "[Getting] Chu was a coup," adds Carolyn Bertozzi, a University of California, Berkeley, chemist, who also directs the Molecular Foundry, DOE's nanofabrication user facility. "He's really been able to inspire people."

Asked why he decided 2 years ago to quit Stanford University and move up the road to LBNL, Chu doesn't hesitate. "I was drying up as a scientist and had nowhere to go," he deadpans before breaking up. In reality, his lab of 13 students and postdocs was thriving. He moved there in 1987 after 9 years at AT&T Bell Labs in New Jersey, where he did the work that won him a share of a Nobel Prize: developing laser-based techniques to cool clusters of atoms to just above absolute zero. He continues to pursue that work, although he's slowly shifting his lab over to work in biophysics. In truth, he says, he was propelled to make the move by his concern over climate change. "It's quite sobering," Chu says of the scale of the task. "This is a problem we have to address, and we have a limited amount of time to do it. What we do in the first half of this century, we will see the consequences for the next 500 to 1000 years."

Chu knows full well that solving this problem is certain to involve government policy and international diplomacy, and it may even depend more on building codes than science. But he seems to thrive on the multifaceted challenge, gliding effortlessly between diverse topics such as how local tax policy in China stops that country from adhering to its own clean-air regulations, whether CO_2 pumped to the ocean floor will stay put, and the potential of architectural innovations to reduce the amount of energy used to heat, cool, and light buildings.

In the end, Chu says, it will take an array of solutions to hold carbon emissions to manageable levels. And they have to start happening fast. Says Chu: "We don't have the option to fail on this one."

–ROBERT F. SERVICE

CREDIT: ROY KALTSCHMIDT/LBNL

CREDIT: CANBERRA TIMES

PROFILE: ANDREW BLAKERS AND KLAUS WEBER

Eureka Moment Puts Sliced Solar Cells on Track

Stuck with conventional technology, a pair of Australian researchers hit upon a way to take silicon cells somewhere new

A train ride through the Scottish countryside to historic Edinburgh would not seem an ideal occasion for flashes of scientific insight. But that setting produced a eureka moment for photovoltaics researchers Andrew Blakers and Klaus Weber. While clattering along in a rocking carriage, the two Australian National University (ANU) scientists conceived a new way of making solar cells that promises to reduce the cost of panels by as much as 75%. Instead of trying to make the wafers used in solar cells thinner to save on expensive crystalline silicon, they start with a thick wafer and divide it into thousands of long, thin slices that they dub SLIVERs.

"The SLIVER cell is a highly original approach, the best way yet suggested of making small, efficient solar cells," says Martin Green, a photovoltaics expert at the University of New South Wales in Sydney, Australia. A pilot plant in Adelaide is now producing SLIVER cells, and full-scale commercialization could be just 2 years away. The SLIVER cost advantage comes from dramatically reducing the use of expensive monocrystalline silicon, the material of choice for solar cells because of its relatively high efficiency in converting solar energy into electricity.

Researchers around the world are racing to find ways to minimize the amount of silicon needed. In the traditional approach, they cut round wafers from a cylindrical silicon ingot and add dopants and coatings to the surface. To reduce silicon use, some groups are exploring exotic materials, others are trying to form thin films of monocrystalline silicon on a substrate. Blakers and Weber were trying to push the limits of conventional silicon technology.

Blakers, 51, has worked on photovoltaics since his student days. He earned his Ph.D. and did his initial postdocs at the University of New South Wales under Green, whose group holds the efficiency record for silicon solar cells. After a stint at the Max Planck Institute for Solid State Physics in Stuttgart, Germany, Blakers set up a silicon photovoltaics research lab at ANU in Canberra. When other researchers began moving beyond silicon, Blakers and his team found they were held back by funding. "We decided to explore what the possibilities were within the constraints we had," says Weber, who joined the lab as a Ph.D. student in 1993 and is now a senior lecturer.

The two were pondering those possibilities on that Scottish train in 2000 en route to a renewable energy conference in Edinburgh. "It was a true collaboration," Blakers says of the moment of inspiration. It took just a couple of months of work back in the lab to produce their first SLIVER cells.

Their technique starts with a silicon wafer 1 or 2 millimeters thick. They etch slots completely through the wafer, leaving silicon strips connected to the edge of the wafer in something that resembles a circular grill. They then dope, metallize, and coat the grill to produce photovoltaic cells, detach the strips from the wafer, and lay them on their sides for encapsulation. Each strip forms a SLIVER cell 50 to 100 millimeters long, 1 or 2 millimeters wide, and 40 to 60 micrometers thick.

Processing wafers this way produces cells that contain just one-tenth the silicon used in normal solar cells with the same surface area. And that will cut the cost of finished modules from the current $4.50 per watt to something approaching $1 per watt, Blakers says. Because the process treats both sides of the SLIVERs, they could be used in applications to capture both direct and reflected light. Spaced SLIVERs in a glass sandwich create a transparent panel that could act as window glass while still generating electricity. And because they are thin, SLIVERs are flexible enough to be used in roll-up solar panels.

Origin Energy, an Australian natural gas and electricity generator and supplier, licensed the SLIVER technology in 2003 and built a pilot plant in Adelaide that, company spokesperson Tony Wood says, has verified that SLIVERs can be produced reliably and in quantity. Wood says Origin is now in discussions with "a small number" of potential partners to scale up production. The commercial plant they plan to announce later this year will likely be in Europe or North America to be closer to markets and to take advantage of government incentives for alternative energy.

Wherever SLIVERs are made, "there are important niche applications for this approach, such as in powering portable equipment," says Green, who is working on a rival technique to reduce silicon usage. But, he adds, "I think a factor of 4 clearly would be an overestimation of any cost advantage." He notes that handling the 2000 to 3000 SLIVERs created from each wafer poses a manufacturing challenge.

Blakers foresees cost gains for the SLIVER technology when it moves to mass production. As for handling the SLIVERs, he says they have devised a simple scheme that works "without robots, machine vision, or pick-and-place equipment." He promises that the proprietary process will be revealed at conferences sometime this year, and it's nothing that couldn't be imagined on a train ride through Scotland. —DENNIS NORMILE

Soft cell. Chapple's work in poplar suggests a cheaper source of ethanol.

PROFILE: CLINT CHAPPLE

How to Make Biofuels Truly Poplar

Clint Chapple has long reveled in the details of plant biochemistry. Now he's finally getting his hands dirty

A master of the biochemical minutiae of plants, Clint Chapple never expected to conduct research requiring a diesel excavator and a greenhouse for growing trees.

But that's where Chapple now finds himself as he explores whether poplar trees could replace corn as a crop for making ethanol for fuel. Over the past 15 years, Chapple, a plant biochemist at Purdue University in West Lafayette, Indiana, has redrawn the map of a key metabolic pathway in plants—work that has won him a solid reputation among academic researchers. Yet plant scientists such as him have long remained second-class citizens in the life sciences funding game. Now, however, with the government embracing biofuels, basic researchers such as Chapple are finding themselves thrust out of their academic cloisters and into the spotlight.

"It's like they say—the right place at the right time," says Chapple, 47. But Chapple's one-time mentor Chris Somerville, a biochemist at Stanford University in Palo Alto, California, says Chapple is being modest. In fact, Somerville says, Chapple has continually exploited unanticipated influences on his career—$70-a-barrel crude oil being only the latest. A fascination with plants' "amazing chemical repertoire" led to graduate work in biochemistry for Chapple, who arrived at Somerville's lab in 1990 with no training in genetics. But he quickly learned enough to create mutants of the model species *Arabidopsis* and develop a promising metabolic method of characterizing polysaccharides in the cell wall (*Science*, 20 August 1993, p. 1032).

A related side project led to a discovery in the pathway that creates lignin, a crucial cell wall stiffener and, behind cellulose, the second most abundant polymer on Earth. That find, a description of the enzymatic step that leads to both straight and branched forms of lignin, helps explain why certain plants break down more easily than others. Over the next 15 years, Chapple uncovered key genetic mutations and steps in the pathway, amounting to a "huge" contribution, says chemist and colleague John Ralph of the University of Wisconsin, Madison.

Lignin is a nuisance to ethanol makers, as it blocks enzymes from getting at the cellulose to break it down to make ethanol. So the U.S. Department of Energy wants Chapple to see how cells modified to make different kinds or amounts of lignin could help poplar compete economically with corn as a feedstock. The first step in the 3-year, $1.4 million project is for Chapple to create the DNA tools with which lignin deposition can be modified. Next, Purdue's Richard Meilan, a transgenic tree expert, will generate transgenic tree saplings, and chemical engineer Michael Ladisch will evaluate the difficulty of recovering cellulose from each. "It dovetails so nicely with the work I've done," Chapple says.

The applied nature of the work is a new twist for Chapple, who says he "always assumed someone else would do it." Experts say the research is not groundbreaking in terms of basic science, but it does carry risks. Chief among them is whether the public will reject transgenic trees. Researchers are buoyed by widespread U.S. acceptance of modified food crops, now common in U.S. grocery products. But remaining European resistance and Greenpeace U.S.A.'s belief that transgenic trees are a "danger in the woods" make supporters such as biochemist Shawn Mansfield of the University of British Columbia in Vancouver, Canada, believe it could be "a hard sell."

One part of Chapple's answer is to find metabolic fingerprints for the transgenic varieties the team grows. That information might enable researchers to find similarly valuable trees in nature and breed them. But that points to another challenge: the "nightmare," as Meilan puts it, of teasing apart the myriad peaks on chromatograms to nail down which part of the fingerprint corresponds to the handful of critical metabolites out of thousands that plants make.

Colleagues say Chapple has the right makeup for the job, and he's included in the project consultations with experts on the societal and ethical impacts of the work. Despite Chapple's limited experience in working with industry partners, for example, Meilan says Chapple wisely addressed concerns of visiting company officials by emphasizing Purdue's experience with government regulators.

Regardless of how effectively Chapple's team overcomes the obstacles to making poplar a legitimate fuel crop, one thing is clear. Plant scientists are being lavished with newfound attention and cash. Just as gratifying, says Chapple, are the shifting attitudes of his department's graduate students, who he says have usually viewed research on plants as less attractive than biomedical pursuits. "I never considered plant sciences before," says Anton Iliuk, a first-year Purdue graduate student in biochemistry. "All of a sudden it feels very exciting."

—ELI KINTISCH

CREDIT: TOM CAMPBELL/PURDUE AGRICULTURAL COMMUNICATIONS

PROFILE: DEBRA ROLISON

Small Thinking, Electrified Froth, And the Beauty of a Fine Mess

An electrochemist seeking better building materials for power-generating devices hymns the virtues of "high-performance disorder"

For someone who works for the U.S. Navy, Debra Rolison doesn't mind rocking the boat. Ask her about fuel cells, and she wastes no time telling you that the way most other chemists build their electrodes—stacking regular crystals into structures tens of micrometers in size—is the wave of the past. The *future*, Rolison predicts, will be both smaller and messier.

"We've lived too long in the age of masonry, the glop-and-slop approach to making things like batteries and fuel-cell electrode structures," she says. On the nanoscale—at which chemistry happens, however, materials are a far cry from tidy brickwork. They're badly mixed, irregularly structured, shot through with holes. Chemists must harness that irregularity, Rolison says, to build better devices from the bottom up.

"She makes people think," says Henry White, an electrochemist at the University of Utah in Salt Lake City. "Even if you don't agree with her, she can really stir things up and make things happen."

What has stirred Rolison, 52, up for the past 36 years is chemistry. She says her passion ignited when she studied the subject in high school and realized that she had "an instinctive understanding of redox reactions." She sailed through college in 3 years ("almost two") and in 1980 received her Ph.D. from the University of North Carolina, Chapel Hill. Uninterested in an academic career but wary of chemical companies' reputation for shunting women scientists into management jobs, she signed on as a research chemist at the Naval Research Laboratory (NRL) in Washington, D.C. She has been there ever since. Now as head of NRL's Advanced Electrochemical Materials Section, Rolison supervises two staff scientists, four postdocs, and usually a couple of undergraduate researchers. "We're small, but we have our fingers in a lot of pies," she says.

Rolison's team first broke the crust of energy research in the mid-1990s with a Defense Advanced Research Projects Agency project to create new high-performance catalysts for methanol-powered fuel cells. Catalysts promote reactions at a fuel cell's two electrodes, the anode, which strips protons and electrons from the cell's fuel (hydrogen or a hydrogen-rich chemical), and the cathode, which splits oxygen molecules from air so the cell can form its main waste product, water. In combination, the reactions create an electrical current that can be harnessed for power. Better catalysts mean more power.

New wave. Rolison says unconventional materials may hold the key to higher-powered fuel cells.

CREDIT: COURTESY OF D. ROLISON

In an early project, Rolison's lab scrutinized nanoscopic bits of platinum-ruthenium, a popular electrocatalyst for direct oxidation of methanol. The material was "a mess," Rolison says, a carpet of ruthenium oxide festooned with water molecules, overlying a platinum-rich core. Yet it was a better catalyst than orderly platinum-ruthenium alloy. "Nature was creating an optimal intersection for adsorption and charge transport," Rolison says. Since then, she has kept an eye out for cases of "high-performance disorder."

That fascination with messiness drew Rolison to aerogels, rigid but incredibly lightweight materials riddled with nanoscopic pores. Rolison and colleagues brew their gels so that the pores (which make up about 80% of the materials) stay connected, creating labyrinths of tunnels through which molecules can pass almost unhindered. A cubic centimeter of aerogel may contain a total surface area approaching 100 square meters—just the sort of "real estate" you need in catalytic electrodes, Rolison says. Her lab and others are working on lining the minute tunnels with electrocatalysts. If they succeed, an aerogel electrode would form a three-dimensional (3D) matrix unlike anything currently in use.

When this was first proposed, it took a lot of criticism as being pure science fiction," says White, who also investigates 3D electrochemical systems. But now, he says, mainstream energy researchers are taking Rolison's ideas seriously. One is Ralph Nuzzo, a chemist who researches fuel cells at the University of Illinois, Urbana-Champaign. If anything, he warms half-jokingly, a 3D energy device risks being *too* effective: "You have to find a way to make it work without being a bomb."

Off the job, Rolison has been known to lob a few metaphorical bombs herself. Since 1990, she has campaigned forcefully on issues affecting women in science, including childcare for university researchers and efforts to close a statistical "gender gap" in the upper ranks of scientific leadership. (Both Rolison and NRL representatives stress that she undertakes these activities on her own time and without the lab's involvement.) In 2000, in an editorial in *Chemical & Engineering News*, she called for the U.S. government to enforce equality in university chemistry faculties by invoking Title IX, a law that withholds federal funding from schools found to discriminate against women. Her take-no-prisoners approach makes some colleagues uneasy, but Rolison says she is "unrepentantly unapologetic" about it. "I've yet to understand the squeamishness by so many scientists and administrators with regard to obeying the law," she says.

When it comes to the laws of nature, Rolison is more patient. It could take years, she acknowledges, for her nanoscale approach to electrochemical devices to bear fruit. "We're at the very bottom of the mountain for fuel cells," she says—but that just means the best scenery is still ahead.

–ROBERT COONTZ

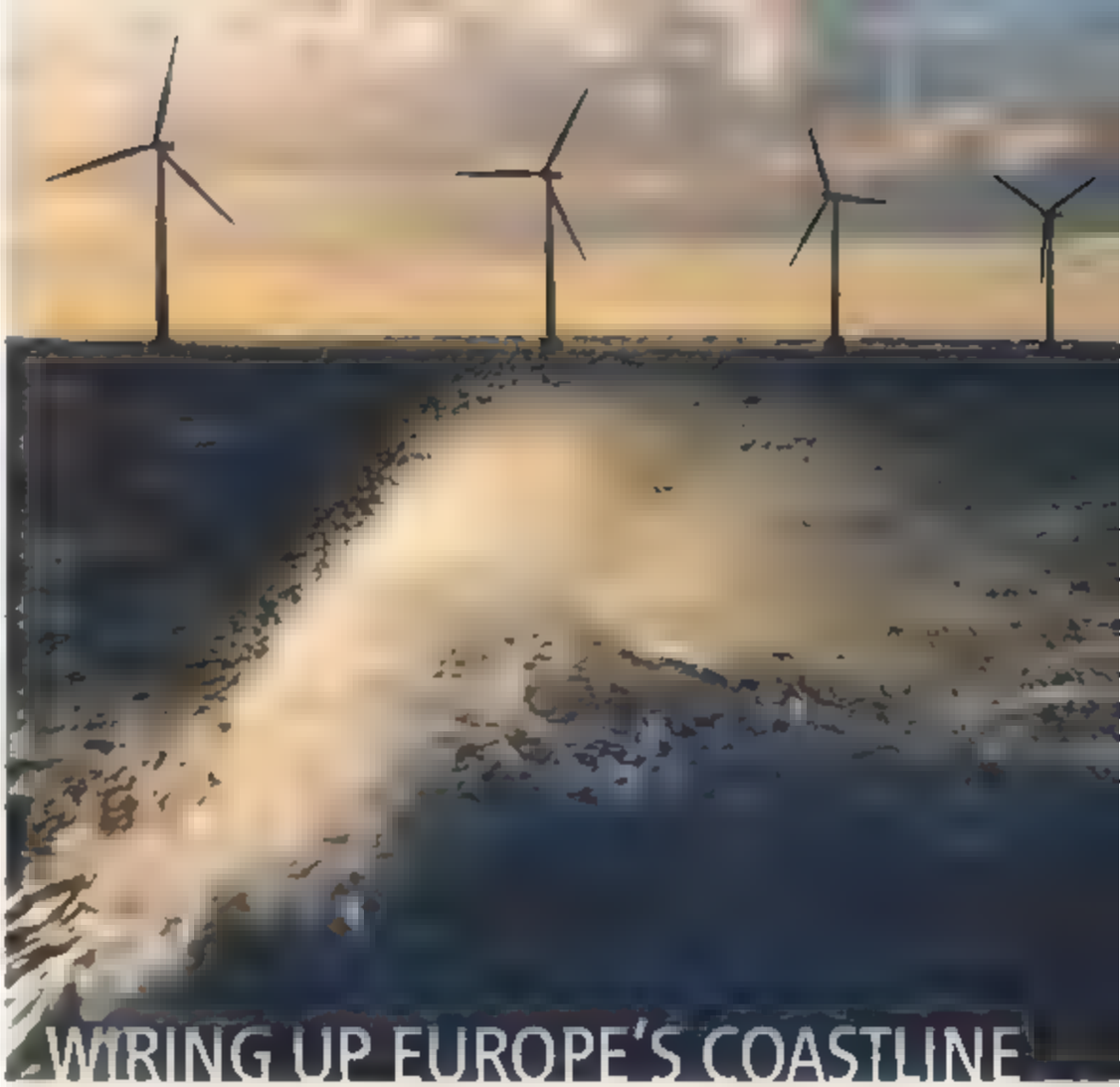
A Fuel for Small Farms

WHEN SHE WAS GROWING UP ON A 12-hectare farm in Monticello, Illinois, Emily Heaton bottle-fed orphan calves and helped move cattle through chutes. Now 28, Heaton is an agronomist with Ceres, a plant and biotechnology company in Thousand Oaks, California, and she sees her job in developing new biofuel feedstocks as a way to save a dying breed: family-owned farms.

Specifically, Heaton focuses on crops such as *Miscanthus*, an incredibly hardy grass that grows to a height of 4 meters and has been widely studied as a biofuel feedstock in Europe but not in the United States. A research project Heaton conducted as an undergraduate at the University of Illinois, Urbana-Champaign, in 2001 showed that the grass could possibly outcompete current American feedstocks of choice, including switchgrass. "It was a very simple project, but it had never been tried," she says of the effort, which the state of Illinois funded in an effort to encourage biofuels.

The work has led to a 40-person project on biomass crops at the university and launched Heaton's career selecting breeds for Ceres to eventually sell. She hopes that work could give small farms the chance to produce crops such as *Miscanthus* for local biorefineries. "There's plenty need for any producer to contribute," says Heaton. She is also trying to create new varieties of feedstocks that can survive tough conditions. "I really get to make new plants," she says of her job. Her father, John Caveny, is also getting in on the act. He maintains a 0.6-hectare stand of *Miscanthus* that he uses to calculate carbon flux for farmers seeking credits for carbon sequestered in plants or soil.

–ELI KINTISCH



WIRING UP EUROPE'S COASTLINE

CRITICS OF WIND POWER OFTEN POINT OUT that when the wind doesn't blow, the current doesn't flow. A Dublin-based wind power company has come up with a novel solution to that problem: a subsea electricity grid linking offshore wind farms all around the coasts of Europe, from the Baltic and North Seas to the Mediterranean in the South. With wind farms far enough apart to be in different weather systems, it will always be windy somewhere around the grid.

Airtricity, a developer and operator of wind farms in and around Ireland and the United Kingdom, including the 520-megawatt Arklow Bank project under construction in the Irish Sea, came up with the idea of a European Offshore Supergrid in 2001. There are now many offshore wind farms built or planned along the coasts of Germany, the Netherlands, the U.K., and other countries. These generally only serve domestic markets, but the Supergrid proposes hooking wind farms to a grid that extends throughout the North Sea, into the Baltic, and then, via the English Channel, the Irish Sea and the Atlantic coast of Ireland. From there, the grid would extend south across the Bay of Biscay and then over the Iberian peninsula into the western Mediterranean.

As a test bed for the technology, Airtricity is proposing to build a huge wind farm in the southern North Sea to supply electricity to the U.K., the Netherlands, and Germany. Dubbed the 10GW Foundation Project, it would consist of 2000 turbines, each capable of generating 5 megawatts, and would cover 3000 square kilometers. Its 10 gigawatts would be enough to power more than 8 million homes. "Good wind resource, shallow water, high demand, countries interconnected—it all comes together in the southern North Sea," says Chris Veal, head of the Supergrid project with Airtricity.

Airtricity is currently seeking support for the project from the three national governments as well as the European Union. Apart from providing zero-carbon sustainable power, the Supergrid has another advantage: It would provide a ready-made way of moving power around the E.U. Most electricity in Europe is generated and distributed by national power companies. There is little trade across borders—something that the E.U. is trying to encourage. "The European Commission and members of the European Parliament all thought it was a good idea. It ticks so many boxes," Veal says. Building offshore is expensive, points out Graeme Cooper of the British Wind Energy Association, but although the economics of the Supergrid might not stack up now, the cost of energy is increasing. "There will be a point when someone will say, 'This is a no-brainer.'"

–DANIEL CLERY

CREDITS: TOP TO BOTTOM: GETTY IMAGES; JINWEI/SPY OF ILLINOIS, URBANA-CHAMPAIGN

PROFILE: DANIEL NOCERA

Hydrogen Economy? Let Sunlight Do the Work

Looking for a clean way to produce hydrogen, Daniel Nocera wants to run a fuel cell backward, powered by sunlight

Hydrogen seems like an ideal fuel. Combine it with oxygen in a fuel cell, and it produces water and electricity, without the noxious pollutants that accompany the burning of most fossil fuels. But it has a dark side. Although there is plenty of hydrogen around, it's bound with other atoms in complex molecules, and it takes large amounts of energy to strip it free. Moreover, most hydrogen today is made from fossil fuels, releasing vast quantities of carbon dioxide in the process. Daniel Nocera is hoping to change that. The Massachusetts Institute of Technology (MIT) chemist is looking for new ways to use sunlight to split water into oxygen and hydrogen. In essence, Nocera is trying to run a fuel cell in reverse. "Why can't we reengineer the fuel cell, backwards?" he asks. "Conceptually, it's very easy."

His quest had a colorful start. Like many fellow Grateful Dead fans with their tie-dyed T-shirts, Nocera, 49, was fascinated by colors when he was growing up—not by the colors themselves, however, but by the processes that lie behind them. "I wanted to understand the colors of materials," Nocera says. And that sparked his interest in chemistry. In between going to some 80 Dead shows, he learned that the colors we see are driven by which frequencies of light are absorbed and reflected by materials. When that light is absorbed, it kicks electrons out of their relaxed state, sending them into a dancing frenzy. As they return to their relative rest, they give off heat. By the time Nocera was a graduate student at the California Institute of Technology in Pasadena, he was wondering how he could design systems to capture electrons excited by sunlight and use them to make fuel. "Right out of the blocks, I was interested in solar energy conversion," Nocera says. Or as Deadheads might put it, finding a way to hold onto the light.

Plants do this by photosynthesis. Two massive protein complexes split water and carbon dioxide and forge new energy-storing bonds in sugar molecules. At the heart of the process is the harnessing of electrons excited by sunlight. Nocera is looking for novel catalysts that will perform some of these tasks more efficiently.

CREDIT: COURTESY OF D. NOCERA

Uphill battle. Nocera hopes to find new catalysts that harness sunlight to make hydrogen fuel.

Five years ago, he and then-graduate student Alan Heyduk designed a ruthenium-based catalyst that uses light energy to strip protons and electrons from an acid and stitch them together to make H_2 molecules. (They're still trying to find a cousin that does the same thing with water rather than an acid.) But that was the easy step. It's even more difficult to grab lone oxygens liberated by splitting water and link them together to form O_2—a step that would be needed to complete the water-splitting reactions and maximize the efficiency of the process. "That requires a certain organization," says Heyduk, now a chemistry professor at the University of California, Irvine. And it's one of the reasons there's been so little progress in recent decades in devising catalysts that turn out O_2.

But there could soon be a new glimmer of progress. Nocera says his group is close to publishing results with a new ruthenium-based catalyst that absorbs light and uses the energy to stitch oxygen atoms together to make O_2. The new catalyst isn't very efficient yet. But it works the way he expected it to, Nocera says, which gives him confidence that they're beginning to understand how to control the motion and bonding of different atoms with their new catalysts. Although Heyduk says he hasn't heard the details of the new catalyst yet, "any system where they can turn water to O_2 is a huge advance because of the difficulty of the problem."

Nocera acknowledges that even this and other advances his group has made are baby steps compared to what's needed for an industrial version of the technology. One problem is that Nocera's catalysts thus far contain ruthenium, a rare and expensive metal. But Nocera is hoping to find cheaper materials that work as well. He recently launched a new project with MIT colleagues to synthesize a multitude of novel metal oxide compounds for testing as possible water-splitting catalysts.

A handful of other groups around the world are engaged in related efforts. One approach pursued by researchers at the National Renewable Energy Laboratory in Golden, Colorado, for example, uses semiconductors to absorb sunlight and create electrical charges that are then used to split water. Although this approach is currently more efficient than Nocera's catalysts, the semiconductors needed are still too costly to be commercially viable. At this point, Nocera argues, all such strategies are worth pursuing. "I'm not sure what the winner will be that is able to make energy without adding extra CO_2 to the atmosphere," Nocera says. "A failing of energy R&D for the last 30 years in the United States has been that it has been treated as an engineering problem with a little 'r' component and a big 'D.' There needs to be an 'R' bigger than the 'D.' There are whole new areas of science and engineering that need to be discovered to solve this problem."

Still, Nocera is convinced that the broad community of researchers now being inspired to find a carbon-neutral source of energy will succeed. "I'll guarantee it," he says. "I think science can deliver a cheap and efficient solution. I believe it deeply."

–ROBERT F. SERVICE

Going green. BP wants some of its future fuels to come from a field, not a well.

PROFILE: STEVEN KOONIN

Guiding an Oil Tanker Into Renewable Waters

Steven Koonin gave up academic life in California for the harsh realities of the energy industry, and he's having a whale of a time

LONDON—Steven Koonin's career took a surprising turn in 2004. After nearly 3 decades as an academic theoretical physicist, including 9 years as provost of the California Institute of Technology (Caltech) in Pasadena, he packed up his southern-California roots and moved to London to become chief scientist of the oil company BP. But his new role is not all about oil. BP was the first major oil company to acknowledge publicly that people may be causing climate change. Now, committed to move "beyond petroleum," it is complementing oil exploration with investment in solar and wind power, clean coal technology, and biofuels. Koonin's job is to help allocate BP's $500-million-a-year research funding, plot its technology strategy, and generally contribute ideas about the energy challenges facing the world. "Sometimes I feel like I've got the most wonderful job in the world," Koonin says.

Speaking in BP's smart headquarters in St. James Square, Koonin, 55, reflects on a career that has also taken him into climate research and science advice. "I like finding out about the way the world works, and ... I like learning new things." He's a quick study. Colleagues and energy-policy experts are impressed by his knowledge of what is essentially a new field to him. "He's risen to the challenge and has enormous influence in the company," says physicist Ernest Moniz of the Massachusetts Institute of Technology (MIT) in Cambridge, a former Department of Energy undersecretary who also sits on an advisory panel for BP. "He's well past the hump on the learning curve," says solar physicist Philip Goode of the New Jersey Institute of Technology in Newark.

Koonin was born and raised in New York City, educated at Caltech and MIT, and joined the Caltech faculty in 1975 specializing in nuclear and many-body theory. Computers were developing rapidly at the time, and Koonin found "lots of things to apply computing methods to." At one point, he says, a representative from IBM offered him a new device called a personal computer to see if he could find uses for it. In the mid-1980s, he got involved in science advice to government, eventually joining JASON, an independent group of scientists that advises the government on science and technology.

Alive to the challenge. Koonin wants to "do biofuels right."

Some work for JASON led to the Earthshine project, an effort to monitor Earth's reflectance by measuring how much earthlight falls on the moon's dark face. "Earthshine helps us to understand the role of clouds in climate," says Goode, Koonin's collaborator.

After he became Caltech's provost in 1995, Koonin oversaw a major overhaul of the institute's biological sciences and expansion of neurosciences, as well as its involvement in projects such as the Thirty Meter Telescope. But in 2003 when a colleague told him BP was looking for a chief scientist, he decided to take the plunge. "I wanted to find out how the private sector works," he says, and energy "was going to be interesting."

Koonin says he spent his first year and a half on the job learning about energy, a process that changed his views. "I was more skeptical about climate change a few years ago. Now I've come round more toward the IPCC's view." (The Intergovernmental Panel on Climate Change has concluded that temperatures are rising in part as a result of human activity.) Like most energy pundits, he sees no silver bullet that will save the planet from climate change, but "some ammo has a bigger caliber than others." He advocates large-scale efforts in carbon sequestration at fossil fuel–burning plants, as well as a new generation of nuclear power stations. And he says it would be "irresponsible" not to investigate other ways to deal with global warming, such as geoengineering (increasing Earth's reflectivity by pumping material into the upper atmosphere).

One area claiming much of his attention is biofuels. Current biofuel efforts, he argues, are strapped onto agricultural food production and are "not optimal." Koonin wants to galvanize geneticists, biotechnologists, agricultural scientists, engineers, and others to "do biofuels right." BP has put up $500 million over 10 years to create an Energy Biosciences Institute (EBI). Five universities are vying for this prize, and the winner should be announced by the time this issue is published.

This initiative has won plaudits. "Koonin thought hard about how to structure the EBI, and it will have a lot of impact," says carbon sequestration expert Robert Socolow of Princeton University. "BP has made a commitment to go big in energy biosciences. I doubt this would have happened without Steve Koonin," says Moniz.

Koonin is pleased with the buzz the EBI has caused. "Plant geneticists are talking to chemists and engineers. ... Researchers are coming alive to the challenge," he says.

—DANIEL CLERY

CREDITS (TOP TO BOTTOM): DAVID PAUL MORRIS/GETTY; COURTESY OF S. KOONIN

PROFILE: TARIQ RAUF

Treading the Nuclear Fuel Cycle Minefield

Tariq Rauf has the unenviable job of making IAEA's international fuel bank work. And the clock is ticking

VIENNA, AUSTRIA—Nuclear weapons capability could spread to as many as 30 more countries in the coming decades if the trade in nuclear fuel continues on its present course, according to Mohamed I. Baradei, director general of the International Atomic Energy Agency (IAEA). But this frightening scenario might be avoided if the "haves" could agree on a better scheme for sharing fuel production with the "have-nots." A number of proposals have now been put forward to allow countries to use nuclear energy without acquiring centrifuges of their own for enriching uranium.

Among these is the establishment of an IAEA-controlled international fuel bank from which all countries could draw. That plan got a boost last year from the U.S.-based Nuclear Threat Initiative (NTI), an independent group backed by U.S. billionaire Warren Buffett. The NTI pledged $50 million to set up the bank, as long as IAEA secures another $100 million, or the equivalent in nuclear fuel, by September 2008.

So far, no one has put money or fuel on the table, and the whole idea remains intensely controversial. "Whether the U.S. would actually place its [nuclear] material under full IAEA control remains to be seen," says Frank von Hippel, a nuclear policy expert at Princeton University. Matthew Bunn, a nonproliferation specialist at Harvard University, says the bank has "a better-than-even chance" of being set up. Others, however, are not happy about the terms being offered. "Forgoing uranium enrichment in order to obtain security of supply is not an acceptable option for many non-nuclear countries," says Jose Goldemberg, a nuclear fuel cycle expert at the University of São Paulo, Brazil.

At the center of this storm sits Tariq Rauf, the 55-year-old head of IAEA's Verification and Security Policy Coordination section and coordinator of the IAEA fuel-bank project. *Science* met with Rauf, a Canadian with Pakistani parents, in his office at IAEA headquarters here in the Austrian capital.

–JOHN BOHANNON

Q: Why is a fuel bank needed?

This whole thing started in the fall of 2003 when our director general [Mohamed I. Baradei] drew attention to the fact that nuclear enrichment and reprocessing technology are in too many hands. Today, there are eight to 10 countries with the capability to enrich uranium and about the same number that can reprocess spent fuel to make plutonium. The question is whether fuel production will be restricted to these countries or whether new ones will enter the market. The issue is that the same technology can be used both for nuclear energy and a nuclear weapons program.

Q: Why would a nuclear hopeful nation want to enroll?

For one thing, enrichment and reprocessing are very expensive activities. Setting up an enrichment plant isn't economical unless you have eight to 10 nuclear power reactors. Some countries do not need so many. This brings up the conundrum: Do you make or do you buy [nuclear fuel]? So the [fuel bank] idea is to have a system whereby [countries] first go to the market to buy fuel, and if they are unable to because of political reasons, then they would come to these assurance of supply mechanisms. It's like if you have bought a ticket from an airline and that airline company goes belly-up. Another airline will honor that ticket.

Q: Besides the political differences, are there technical challenges?

The challenge is to have it be multinational without a transfer of technology. For example, if you had six countries taking part, the enrichment technology might be coming from the Europeans, and [they] would run the technology. The other countries are part of the management and operational side.

Q: So scientists and technicians from every country would not be involved?

They could be involved, but they wouldn't all be sitting in the [enrichment] cascade halls. None of these multinational schemes envisions the expertise of enrichment or reprocessing being transferred to countries that don't have these technologies already. It's as if you bought shares in a company like Toyota. You're interested in the product, which in this case is the enriched uranium coming out. You really don't need to know how the production line works.

Q: What role is IAEA likely to play?

One of the criticisms is that this is a grand plan from the IAEA to expand and be a supercontrolling agency. … But we don't want to set up an empire. If we do set up an IAEA fuel bank, we would likely contract it out to industry.

Q: What might a nuclear renaissance mean for nuclear science?

The intake of people studying nuclear science in universities has been declining, which has made a smaller and smaller pool of nuclear-educated people available. That has also made life difficult for [the IAEA] because we need inspectors and other staff with nuclear expertise. Many of us hope that a nuclear renaissance will mean that nuclear will no longer be associated with being unsafe and that this will encourage students. The world certainly needs more people with a nuclear science education.

CREDIT: IAEA

Norway: A Nuclear Demonstration Project?

EGIL LILLESTØL IS A MAN WITH A RATHER unusual mission. He wants his homeland of Norway to take the lead in developing a new form of nuclear power. Norway is Europe's largest petroleum exporter, from its North Sea oil and gas fields, and Lillestøl, a physicist at the University of Bergen, believes the country needs to do something about its carbon emissions. Norway has little experience with nuclear power but has one of the world's largest reserves of thorium. Lillestøl says Norway should pioneer a new, inherently safe form of nuclear reactor called an energy amplifier that runs on thorium. "It would be a good thing to have other [options] to stand on," Lillestøl says.

Carlo Rubbia, a Nobelist and former director general of Europe's particle physics lab CERN, championed the idea of the energy amplifier in the 1990s, and CERN researchers developed a design and tested some of the key ideas. A conventional fission reactor holds enough fissile material for a nuclear chain reaction to take place, neutron-absorbing rods ensure that the reaction doesn't run out of control, although this always remains a risk. The energy amplifier doesn't have enough fissile material to sustain a chain reaction. Instead, an accelerator fires high-energy particles into the fuel, prompting a cascade of fission reactions and producing heat. The amount of heat is proportional to the intensity of the beam, and the accelerator can be designed so that the amplifier can never overheat. Although the amount of waste produced is expected to be low, particle accelerators aren't cheap, and one with the necessary power has never been built.

Lillestøl wants Norway to pioneer this form of energy by funding and hosting a prototype—at a cost of about €550 million—and has made it a personal crusade to win over the Norwegian public and government. Lillestøl says he makes two or three presentations a week. Although the government is wary of nuclear power after a debate in the national assembly, the energy minister called for an in-depth study. Norway is in a unique position to undertake such an enterprise because it has been squirreling away oil revenue and has now amassed a fund of some $250 billion.

CERN's Jean-Pierre Revol, who worked on the energy amplifier at CERN, says that Lillestøl has made "a lot of political progress" in Norway. Renewed interest in nuclear power is generating curiosity about this technology, Revol says. "If it starts to fly, everyone will want to be part of it."

—DANIEL CLERY

CREDIT: DAVID FAIMAN

PHOTOVOLTAICS IN FOCUS

NEGEV DESERT, ISRAEL—What looks like an upside-down umbrella made of mirrors is the future of renewable energy—at least according to its creator, David Faiman, a physicist who directs the Ben Gurion National Solar Energy Center here. Photovoltaic cells have been around for decades, but they've never been competitive with fossil fuels. Faiman claims to have found a way to slash the price. "This technology is a real contender as a solution to the world's energy problem," says physicist Robert McConnell of the National Renewable Energy Laboratory in Golden, Colorado.

The secret ingredient is perched at the focus of his 10-ton reflector: a square grid 10 centimeters across called a concentrator photovoltaic (CPV) cell. Instead of spreading solar panels across a broad area to capture photons, Faiman uses a reflector to concentrate the light 1000 times onto a small target. Traditional silicon-based solar cells can't handle the heat, but a gallium arsenide–based cell developed by a team at the Fraunhofer Institute for Solar Energy Systems in Freiburg, Germany, "actually works far more efficiently at higher temperatures," says Faiman. By using a concentrating reflector, the system makes best use of its most expensive component, the CPV cell, which Faiman estimates can reach 40% efficiency at converting sunlight to electricity. With this system, Faiman believes he can build a power plant for less than the magic number of $1000 per kilowatt of electrical capacity. "Getting the price that low is feasible, but only on a large scale," says McConnell, "and there's a long way to go from this stage."

Large is exactly the scale Faiman is thinking along: spreading 20 000 CPV cells over an area of 12 square kilometers to generate 1 gigawatt. With mass-produced CPV cells, Faiman's estimates the cost at $1 billion. "Considering the savings, the system can pay for itself within 2 decades," he says. The team is hoping to make it happen sooner by increasing the efficiency of the CPV cell, for example by adding extra layers of solar cells that capture a broader range of the wavelengths.

—JOHN BOHANNON

PROFILE: JAY KEASLING

Rethinking Mother Nature's Choices

Jay Keasling believes ethanol is a poor biofuel. So he's going to get microbes to make something better

There are good reasons gasoline has been king of the road for more than a century. It packs lots of energy for its volume, and it's stable, transportable, and noncorrosive. Alternative fuels, like ethanol, have a hard act to follow. Unlike gasoline, which is a collection of medium-length hydrocarbons, ethanol is a short-chain alcohol, sporting a less energetic carbon-oxygen bond alongside some of the power-packing carbon-hydrogen bonds of gasoline. It delivers about 30% less energy by volume than gasoline. It's still useful as a gasoline additive for cars, but it can't do it all. "We're not going to be putting much ethanol in planes," says Jay Keasling, a chemical engineer at the University of California, Berkeley, and the Lawrence Berkeley National Laboratory. "We probably wouldn't choose ethanol as a fuel if nature didn't give it to us."

Interior designer. Keasling is betting that synthetic biology will allow him to make microbes that turn out transportation fuel.

But, says Keasling, "we don't have to accept what nature has given us." Keasling is trying to force nature to reconsider her options, and his tool of choice is synthetic biology, in which researchers add and remove entire biosynthetic pathways to the genomes of organisms to get them to produce better drugs, materials, or—perhaps, one day—fuels. Keasling is one of the discipline's leading advocates and practitioners. And he's betting that by inserting novel pathways into yeast and other microbes, they could produce long oil-like hydrocarbons, or shorter versions called alkanes that can be stitched together by means of simple chemical processing.

No one has been able to do this yet, but Keasling and his students are working on the problem. "Jay has a good vision," says Chris Somerville, a plant biochemist at Stanford University in Palo Alto, California, and co-founder of LS9, a synthetic biology start-up in San Carlos, California. "In about 15 years, we will be making biofuels other than ethanol, and synthetic biology will be what makes that possible," he predicts.

Brazil derives about one-third of its transportation fuel from ethanol made from sugar cane. But in the cooler U.S. climate, corn is king. Keasling, 42, grew up on his family's 200-hectare corn and soybean farm in Harvard, Nebraska, and he still looks like he belongs there, with his close-cropped hair and barrel-like chest. He guffaws when asked whether he considered staying to become the fifth generation of his family to work the farm. "No. The stork dropped me in the wrong place," he quips. Keasling slipped away to the University of Nebraska, Lincoln, then on to a Ph.D. at the University of Michigan and a postdoc with Stanford University Nobel laureate Arthur Kornberg. "It's really hard to make a profit" growing corn, Keasling says, explaining that the price per bushel has remained largely unchanged since the late 1940s. That's prompted farmers to continually try to improve their yields, which in turn has produced a glut of corn. That glut, and high oil prices, has suddenly made corn ethanol more economically viable. "You get off the plane in Nebraska, and it smells like a brewery," he says.

In his effort to beat out ethanol, Keasling won't be starting from scratch. He is already well known for producing an antimalarial drug called artemisinin using synthetic biology. The molecule, a complex double ring structure, is normally derived from a plant. But doing so is expensive. So, over the past 6 years, Keasling's group has inserted more than a dozen genes into *Escherichia coli* and yeast, setting up a molecular assembly line that allows the bugs to churn out the drug. The strategy not only works, but improvements over the years have also boosted the artemisinin output 10-million-fold. Keasling hopes further improvements will make the drug cheap enough to be widely affordable in developing countries, potentially saving millions of lives. Keasling's feats with artemisinin earned him the title of *Discover* magazine's Scientist of the Year in 2006.

Now, Keasling wants to go one better, swapping out some of the ring-building enzymes he introduced into microbes in favor of others that can construct linear hydrocarbons. "We built this platform to produce this hydrocarbon [artemisinin]," Keasling says. "We can remove a few of the genes to take out artemisinin and put in a different hydrocarbon. It's that simple."

At least on paper. Even if the microbes can produce it, Keasling acknowledges that won't be the end of the story. Despite the impressive improvements at getting bugs to churn out an antimalarial compound, they still typically collect only a couple of dozen grams per liter of it at the end of fermentation. That could be a big problem for a transportation fuel that has to be not only cheap but also wildly abundant. "This is massive. This would require a scale-up of fermentation never seen before."

As a result, Keasling argues that it's important for synthetic biology to push other efforts to create alternative fuels. One such idea is to reengineer plants' intertwined networks of lignin and cellulose to make it easier for microbes to break them down and convert them to ethanol. Another is to boost the ability of microbes to make and tolerate ethanol. "This is an exciting time to be in the energy research area," Keasling says. "The biology is just getting to the point where it has something to offer."

—ROBERT F. SERVICE

CREDIT: ROY KAITSCHUCK/LBNL

PROFILE: RICHARD MARTINEAU

Former Marine Seeks a Model EMPRESS

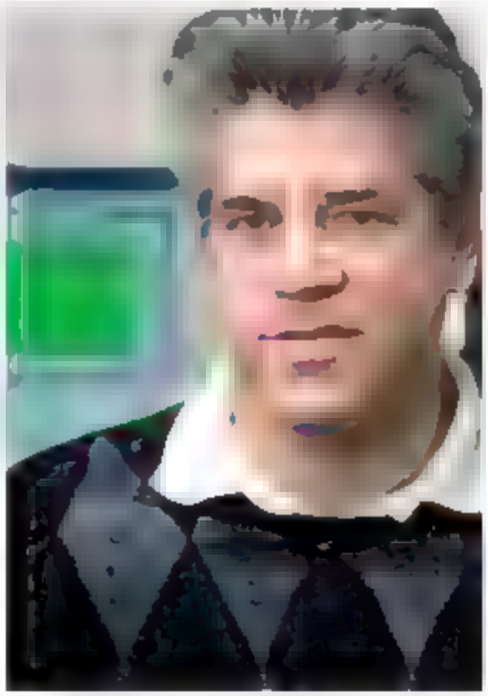
To test nuclear reactor designs, computational physicist Rich Martineau is creating a sophisticated new simulation from the ground up

IDAHO FALLS, IDAHO—Whether he's chewing tobacco or driving his battered '78 Chevy pickup, Rich Martineau stands out in a crowd. The computational physicist aggressively protects his status as an outsider, keeping the uninvited out of his cubicle at the Idaho National Laboratory (INL) in Idaho Falls with a black retractable shade. And the tagline on e-mails from the former U.S. Marine and president of a local gun club reads: "Rome did not create a great empire by having meetings. They did it by killing all those who opposed them."

But in September, Martineau arranged and hosted a decidedly communal gathering, a 2-day, invitation-only conclave on modeling nuclear reactors. Among the 44 participants were experts in applied math, supercomputing, and nuclear engineering, as well as top modeling researchers from the aerospace and nuclear weapons fields. "There was excitement in the air," says Columbia University computational mathematician David Keyes, who adds that the quality of assembled academic firepower "surprised" him. "The Marine image he cultivates is tempered by the realization that you need consensus."

What prompted Martineau, 47, to soften his independent nature is his desire to build a sophisticated digital model of what goes on inside a nuclear reactor. Such a model would be "definitely needed" to jump-start the U.S. nuclear energy industry and a major improvement over much older models currently used by nuclear engineers, says Vincent Chan of General Atomics in San Diego, California.

The goal of the new reactor model Martineau is developing at INL is to allow engineers, government regulators, or plant operators to test novel design concepts using supercomputers instead of costly prototypes. The older models mostly depict antiquated water-cooled, low-temperature uranium reactors, whereas Martineau wants to help industry simulate more innovative designs, including gas-cooled reactors that work at higher temperatures and utilize different fuels such as recycled nuclear waste. "You have to know where your uncertainties come from so that you know [which design decision] matters," says INL nuclear research chief Phillip Finck. More elaborate virtual reactors could also simulate accident scenarios in order to train workers—what INL modeler Glen Hansen calls the "Xbox concept" after the popular video game system.



New build. Martineau wants to "start at the bottom" of nuclear modeling.

Existing reactor computer models haven't been overhauled much since the heyday of the U.S. nuclear enterprise in the 1970s and 1980s. Back then, computational physicists devised simulation programs, including one dubbed RELAP that tried to depict conditions within nuclear reactors under normal and accident scenarios. "Nuclear engineering was, like aerospace, at the cutting edge [in modeling] at the time," says Keyes. But after what Martineau calls "Jane Fonda and Three Mile Island," government research funds dried up. Martineau's work in the 1990s modeled flames and floods, not nuclear reactors. As a result, whereas airplane builders now have fluid dynamics models so robust that they can build whole jets on computers and use fewer wind-tunnel tests, nuclear engineers still depend on crude, 25-year-old computer programs.

Three years ago, Martineau proposed developing a new computational model for reactors to INL manager Kathryn McCarthy. With scant money available from Washington, D.C., McCarthy helped persuade INL chiefs to provide some initial funds for the effort. As a wry sign of thanks, Martineau has dubbed the emerging model Enhanced Multi-Physics Reactor Simulation Suite (EMPRESS). "Kathy had the confidence in me when no one else did," he says of his boss. "After 15 years of stagnating, I exploded into this thing."

Martineau spent the summer of 2004 at Argonne National Laboratory learning about the powerful parallel computing techniques the new reactor model would need. He also began assembling a team. Along with two INL modelers with backgrounds in nuclear engineering, the effort has attracted Dana Knoll and Hansen, both accomplished computational physicists recruited from the strong modeling corps at Los Alamos National Laboratory, a nuclear weapons lab.

The team's goal, as Martineau sees it, is to "start at the bottom," using fundamental physics principles to overhaul the outdated models. RELAP, for example, simulates the heat and neutron behavior of a reactor but only in one dimension and for idealized vessel shapes. Martineau wants EMPRESS to work in three dimensions, incorporating the realistic geometry of a gas-cooled nuclear power plant and avoiding the fudge factors RELAP uses to approximate hard-to-model areas such as gas-vessel boundaries.

Lab officials are realistic about the scope of the project and its prospects of winning federal funding, especially as it's unclear how the new Democrat-controlled Congress will be disposed toward advanced nuclear energy work. "We don't have infinite amount of time and infinite amount of resources to chase down every single thing to the molecular scale," says INL administrator Kemal Pasamehmetoglu.

Still, Martineau's sanguine about the future of EMPRESS. He once dreamt of flying jets in the Marines and now sees his duty as helping expand nuclear power to reduce U.S. dependency on foreign oil. "God, country, and Corps," says Martineau of his allegiances. "I've always wanted to do this."

ELI KINTISCH

CREDIT: CHRIS MORGAN/INL

Options. With student Chris Barrett and others, Dumesic seeks ways to refine carbohydrates.

CREDIT: A. CHO/SCIENCE

PROFILE: JAMES DUMESIC

Catalyzing the Emergence of a Practical Biorefinery

Most are betting on biology to convert carbohydrates into biofuels. Jim Dumesic thinks catalysis is the key

The United States is counting on biofuels to reduce reliance on imported petroleum and to cut carbon emissions from vehicles. But most cars won't run on corn oil, so scientists must find ways to convert plant matter into practical fuels. Much hope—and hype—centers on harnessing microbes and enzymes to convert biomass to ethanol. But James Dumesic, a chemical engineer at the University of Wisconsin, Madison, is blazing another trail. An expert in catalysis, Dumesic is searching for a philosopher's stone to turn sugar water into fuels and higher value chemicals.

Dumesic, 57, is one of a growing number of catalysis experts who are using their skills to convert so-called biorenewables, such as sugars and other carbohydrates, into hydrogen, liquid fuels, and precursors for plastics. Those efforts may be crucial for transforming biorefining from a grand ambition to an economically viable reality.

As does a petroleum refinery, a biorefinery will strive to crank out a variety of fuels and chemicals to maximize profit, says Todd Werpy, a chemist at Pacific Northwest National Laboratory (PNNL) in Hanford, Washington, and "catalysis is going to be critical for both fuels and chemicals in the long run." For his part, Dumesic says he is providing industry with options "so that when carbohydrates become more readily available, we can take them different ways."

As raw materials for fuels, petroleum and carbohydrates lie on opposite ends of a chemical spectrum. The hydrocarbons in petroleum are molecules with few functional groups—add-ons such as an oxygen bonded to a hydrogen—a hydroxyl group—that provide ready sites for reactions. So the components of petroleum are relatively stable and require little modification to produce gasoline and other fuels. In contrast, carbohydrates bristle with functional groups that make the compounds less stable and less capable of withstanding the high temperatures used to process petrochemicals. Those functional groups must be stripped away to transform the carbohydrates into energy-packed hydrocarbons.

To pare down carbohydrates, Dumesic employs heterogeneous catalysis, in which the catalyst resides on the surface of a solid support such as powdered aluminum oxide. Ironically, the techniques of heterogeneous catalysis were honed on petroleum chemistry, says Dumesic. "Over the past 50 years, people have learned a lot about how to make, characterize, and test catalysts, and all of that carries over directly" to carbohydrates.

In 2002, Dumesic and colleagues used platinum to catalyze the production of hydrogen gas from a solution of carbohydrate in water. Dubbed aqueous phase reforming (APR), the low-temperature process requires less energy than traditional steam reforming, in which water vapor interacts with methane. The advance links hydrogen technologies and biorenewables, says James Jackson, a chemist at Michigan State University in East Lansing. "There is no green source of hydrogen because it comes primarily from steam reforming of natural gas," he says. "This would be making it out of biomass sources."

To commercialize the APR technology, Dumesic and collaborator Randy Cortright spun off a company called Virent Energy Systems. The 4-year-old start-up has developed a unit that converts glycerol—a cheap byproduct of biodiesel production—into hydrogen and methane and burns the mixture to drive a 10-kilowatt generator. Virent sold the device to the local utility company, and for the past year it has pumped electricity into the grid.

Virent hasn't turned a profit yet, however. The company is still lining up investors, such as Cargill and Honda, and working from grants, including $2 million from the U.S. departments of Agriculture and Energy. "Don't congratulate me, because we don't have money from product sales coming through the door," Cortright says.

When fed larger sugar molecules, APR also tended to produce methane and other hydrocarbons called alkanes. So in 2004, Dumesic and colleagues changed the catalyst to maximize the production of alkanes instead of hydrogen, potentially opening a new route to liquid fuels. The researchers also added a step to link shorter alkanes into longer ones like those in diesel fuel.

Most recently, Dumesic and his team have found a way to convert the sugar fructose into a compound known as hydroxymethylfurfural (HMF). HMF might serve as a feedstock for making fuels, but it could have even greater potential as a "platform chemical" to produce polymers and other higher-value substances, Dumesic says. PNNL's Werpy also sees the allure in chemicals. "We're focusing much more on chemicals than on fuels because of the bigger margins and opportunities for pulling the field along," he says.

The catalysis of biorenewables is still in its infancy, which makes the field exciting, Dumesic says. "I imagine that there was a period of time in the early days of petroleum chemistry when new things happened all the time, and I kind of have that feeling now as we move into this other area," he says. And this much seems sure: No matter how technologies for biofuels and biorefining evolve, catalysis is sure to be an important part of the mix.

–ADRIAN CHO

PERSPECTIVE

Don't Forget Long-Term Fundamental Research in Energy

George M. Whitesides[1]* and George W. Crabtree[2]

Achieving a fundamental understanding of the phenomena that will underpin both global stewardship and future technologies in energy calls for a thoughtful balance between large-scale immediate solutions using existing technology and the fundamental research needed to provide better solutions in the 50-year period.

Energy and climate change are now preoccupations shared by science, engineering, and society. There is a range of views on energy and almost religious levels of advocacy for particular technologies. There is also surprisingly broad (although not universal) agreement that there is no single solution to the dual problems of meeting future demands for energy and managing the environmental consequences of energy production. Whatever strategy emerges will be a quilt made up of patches representing almost every imaginable technology.

The energy problem is often phrased in terms of developing a strategy that roughly doubles the global production of energy by 2050 (from 13 to about 30 terawatts) (Fig. 1) (*1*–*9*). The problem of climate change includes two especially important components: (i) understanding the relationship between the climate and the chemistry of the atmosphere and oceans and (ii) predicting the impact of climate control on strategies for energy production. Because atmospheric CO_2 is the dominant greenhouse gas, and because coal is the carbon-rich fossil fuel whose use can most readily be expanded (especially in the rapidly growing economy of China), understanding the linkage between coal and climate is particularly important (*6*).

There is a pervasive sense that "We must do something *soon*." This urgency may be justified, but we must also remember that the problems of providing energy and maintaining the environment are not about to go away, no matter how hard we try using current technologies. In the rush to do something—to find technological solutions to global-scale problems—we should not forget that we must ultimately *understand* them, if we are to find the most effective, sustainable solutions. Fundamental research in science and engineering is important. Understanding phenomena relevant to energy and the environment leads to new technologies and to the ability to control the economic and environmental outcomes of their applications (*7*).

The cost of large technology demonstration projects is enormous and the time to develop them decades, and it is easy to overlook the fundamental research that nourishes them. Today, we have a growing thicket of energy and environmental problems and great enthusiasm for solving them quickly. In fact, 50 years from now, most of these problems (and more) will still remain unsolved. Developing the best patches we can for the immediate problems is one approach. Understanding the underlying problems is another, and one that is at least as important, much less expensive, and perhaps ultimately time-saving. Energy and climate are problems that will extend over decades or centuries, and the unimaginable technologies of 100 years from now will rest on fundamental research that must start now.

What follows is a sketch of nine representative long-term problems in research that are vital to the development of future technology for energy. We emphasize that this list is personal and idiosyncratic; it tends to emphasize materials. Others might select differently, although most lists would probably have areas of broad overlap.

The Oxygen Electrode Problem

A hydrogen fuel cell operates by extracting electrons from H_2 and transferring them through an external circuit to O_2 (to generate H_2O). If the H_2 is generated electrochemically, the reverse reactions take place. In either event, the transfer of electrons from H_2O to one electrode and to O_2 from another are slow reactions and lower the efficiency of practical fuel cells (considering the free energy of the reactions involved). The slow rates of interconversion of $4\ e^- + O_2 + 4\ H^+$ and $2\ H_2O$ exemplify a broader class of reactions in which a single process requires the transfer of multiple electrons. Understanding these reactions and finding strategies for circumventing their limitations are important in developing new, more practical procedures for reactions ranging from the electrochemical production of H_2 and the use of O_2 in fuel cells to the reduction of nitrogen to ammonia.

Catalysis by Design

Many of the reactions that occur in the production of energy involve catalysis: the full set used in the processing of crude oil to fuels; all of the biological reactions involved in photosynthesis, in fixing CO_2, and in biodegradation; the hydration of CO_2 to carbonate ion; the movement of electrons in batteries; the operation of fuel cells; the cleanup of exhaust gas from internal combustion engines; and many others. Given the enormous importance of catalysis in the production and storage of energy, in the production of petrochemicals and the materials derived from them, and in all biological and most geochemical processes, it is astonishing (and a little disheartening) how little is known of the fundamentals of catalysis: how catalysts operate, how to control them, and especially how to generate new ones. Catalysis by design has periodically been embraced as a grand challenge, and periodically abandoned as too difficult, but nanoscience and surface science offer new approaches to this problem. The fundamental study of catalysis must be reanimated across the full spectrum of processes involved in energy and the environment.

Transport of Charge and Excitation

Photoexcitation of the semiconductor or dye component of a solar cell creates an exciton, a separated but associated hole and electron (*4*, *5*). To generate current, the electron must move to one electrode, and the hole to the other, before they combine. These processes are inefficient in materials that might make inexpensive photocells: defective, polycrystalline, disordered, or quantum-dot semiconductors (whether inorganic or organic). Understanding them and circumventing their deficiencies is one key to cost-effective solar cells.

Chemistry of CO_2

CO_2 is a key molecule in global warming (*6*), in chemical and biological fuel production, and in fuel use. We must know everything possible about its physical and chemical interactions. Important topics include new uses of CO_2 in large-scale chemistry (where it has the attractive feature that it has negative cost), new chemical reactions of CO_2, the movement and reactions of CO_2 in the earth, the role of CO_2 in determining the behavior of the atmosphere and oceans, and the chemistry and properties of CO_2 at high pressures. For decades, there has been little research, whether fundamental or exploratory, in this area: it was considered a solved problem.

Improving on Photosynthesis

The process of uptake and fixation of CO_2 in biological photosynthesis is not an optimized

[1]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. [2]Materials Science Division, Argonne National Laboratory, Argonne, IL 60439, USA.

*To whom correspondence should be addressed. E-mail: gwhitesides@gmwgroup.harvard.edu

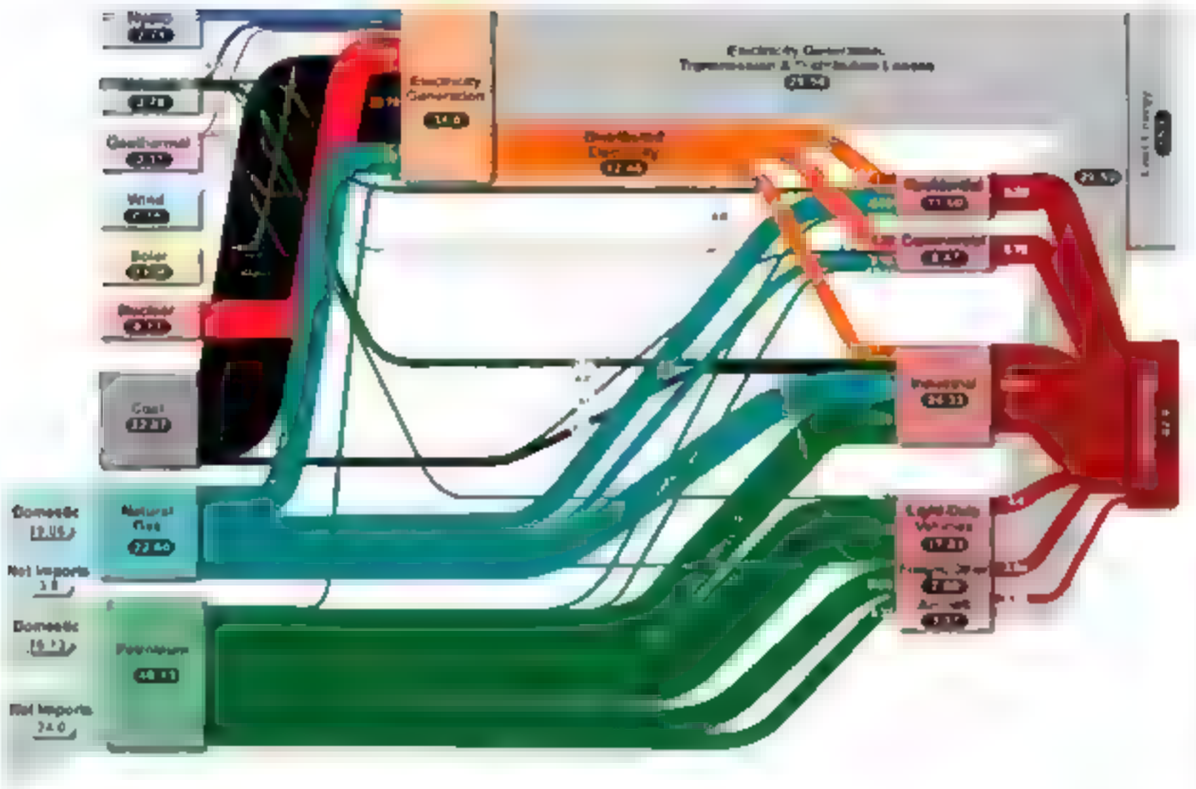


Fig. 1. The complex system of energy flows in the United States in 2005 (2). More than half of the energy produced is wasted. Units are in quads, 1 quad = 10^{15} British thermal units = 1.055 exajoules. (Figure prepared by Lawrence Livermore National Laboratory, University of California, and the U.S. Department of Energy.)

marvel. It is fairly inefficient thermodynamically, and several key reactions (for example, the key step involving the reaction of CO_2 with ribulose diphosphate (with its competitive reaction with O_2) are notoriously poor yield. These inefficiencies are an opportunity. Photosynthesis has immense appeal for the closed-cycle capture of energy from the Sun in forms that are useful as fuels (4). The prospects of re-engineering biological photosynthesis for greater efficiency, maximizing metabolic flux through specific biosynthetic pathways, growing plants in regimes of temperature or salinity where they normally do not flourish, and directly producing fuels such as H_2, CH_4, or alcohols are all alluring ones, but ones that will require decades of imaginative research to realize. The prospect of non-biological photosynthesis (where "photosynthesis" might include bio-inspired physical and chemical reactions or more straightforward photochemical or photothermal processes that generate fuels or store energy) also warrants new research.

Complex Systems

Understanding energy and the environment analytically poses a series of problems that we presently have neither the mathematical tools nor the data to solve. Most global systems are "complex" in physicists' definition of the word. They comprise many components, with many degrees of freedom, usually interacting non-linearly. These systems are the natural home of the surprises often referred to as emergent behavior. Our difficulty in understanding and modeling these systems leads to uncertainties that cloud most discussions of energy and the environment and of the costs and impacts of almost any technology (5). What really is the cost of a kilowatt produced by silicon solar cells? How important will the burning of coal be to global warming? What, in detail, are the global sources and sinks for carbon and how do they interact? What can one say about the impact of technologies for generating nuclear power on the potential for proliferation of nuclear weapons? Development of the theory of complex systems to the point where it gives reliable results (or at least results whose reliability can be quantified) remains a key enabling capability, and is probably the best way of minimizing the potential miseries of the law of unintended consequences.

The Efficiency of Energy Use

Increasing the efficiency of energy conversion and storage is a major opportunity. Many of our standard energy conversion routes are far from their Carnot or exergy limits. Electricity production with the present mix of fuels is only 37% efficient on average; the typical automobile engine is perhaps 25% efficient; and an incandescent bulb is only 5% efficient for producing visible light. Increasing efficiency requires understanding the fundamental phenomena of existing and alternative energy conversions. Solid-state lighting, for example, can achieve efficiencies of 50% or more, provided that we understand the mechanisms controlling the conversion of electronic energy to photons. New understanding of mechanisms of friction, wear, and corrosion also provides new strategies for reducing losses.

The Chemistry of Small Molecules

The chemistry of small molecules dominates many aspects of energy and climate. H_2O, H_2, O_2, CO_2, CO (for Fischer-Tropsch chemistry), NO_x, O_3, NH_3, SO_2, CH_4, CH_3OH, HCl, and others are all vitally important components of these discussions. There remains a wide range of information about these molecules and their combinations that is needed to understand the complex systems of which they are a part.

New Ideas: Separating Wheat from Chaff

The spectrum of ideas for dealing with problems of energy and global stewardship is not complete, based just on what we now know. We need new ideas, and we need to know which of the current smorgasbord of unexplored and unproved ideas will work (9). Developing affordable technologies for removing carbon from the atmosphere (for example by growing biomass and converting it to a stable form of carbon) must be explored now, if they are to be options in the future. Changing the albedo of Earth, stimulating photosynthesis in the oceans by the addition of essential trace elements such as iron, developing new nuclear power cycles, a hydrogen economy, new methods for separating gases (such as CO_2 from air) and liquids, room-temperature superconductivity to carry electric power without loss, biological H_2 production, new concepts in batteries, and nuclear fusion—all must be explored fundamentally and realistically.

These problems all require long-term, patient investment in fundamental research to yield new and validated ideas. These problems are also, in some cases, sufficiently technical that their importance is most obvious to specialists. The oxygen electrode (as one example) might seem an exotic problem in science, but it is hard to

believe that a hydrogen economy that used electrolysis to generate H_2 and O_2 from water, and a fuel cell to convert H_2 and O_2 back to water and electrons, could make a substantial contribution to global energy without a much-improved oxygen electrode. The identification of this problem is not in any sense new. The redox chemistry of oxygen has been a subject of active interest (but limited success) for decades. We simply need new ideas.

Another reason to work on these big problems is that they will attract the most talented young people. Over the past 30 years, the National Institutes of Health has used stable and generous support to recruit and build a very effective community of biomedical scientists. Solving the problems of energy and global stewardship will require the same patient, flexible, and broadly based investment, if society believes that the problems in these areas are sufficiently important to provide a life's work for its most talented young people.

References and Notes

1. President's Council of Advisors on Science and Technology (PCAST), *The Energy Imperative: Technology and the Role of Emerging Companies* (2006), www.ostp.gov/PCAST/pcast.html
2. *World Energy Outlook 2004* (International Energy Agency, Paris, 2004), www.worldenergyoutlook.org
3. *Basic Research Needs to Assure a Secure Energy Future*, J. Stringer, L. Horton, Chairs [workshop report, U.S. Department of Energy (DOE) Office of Basic Energy Sciences, 2003], www.sc.doe.gov/bes/reports/abstracts.html#SEC
4. *Basic Research Needs for Solar Energy Utilization*, N. S. Lewis, G. W. Crabtree, Chairs [workshop report, DOE Office of Basic Energy Sciences, 2005], www.sc.doe.gov/bes/reports/abstracts.html#SEU
5. Systems and life-cycle energy Technology Analyses (National Renewable Energy Laboratory), www.nrel.gov/analysis/tech_analysis.html
6. See discussions of global climate science from the National Center for Atmospheric Research, www.ucar.edu/research/climate
7. J. M. Deutch, E. R. Lester, *Making Technology Work: Applications in Energy and the Environment* (Cambridge Univ. Press, New York, 2004)
8. N. S. Lewis, D. G. Nocera, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15729 (2006).
9. M. S. Dresselhaus, I. L. Thomas, *Nature* **414**, 332 (2001)

10.1126/science.1140362

PERSPECTIVE

Toward Cost-Effective Solar Energy Use

Nathan S. Lewis

At present, solar energy conversion technologies face cost and scalability hurdles in the technologies required for a complete energy system. To provide a truly widespread primary energy source, solar energy must be captured, converted, and stored in a cost-effective fashion. New developments in nanotechnology, biotechnology, and the materials and physical sciences may enable step-change approaches to cost-effective, globally scalable systems for solar energy use.

More energy from sunlight strikes Earth in 1 hour than all of the energy consumed by humans in an entire year. In fact, the solar energy resource dwarfs all other renewable and fossil-based energy resources combined (*1*). With increasing attention toward carbon-neutral energy production, solar electricity—or photovoltaic (PV) technology—is receiving heightened attention as a potentially widespread approach to sustainable energy production. The global solar electricity market is currently more than $10 billion/year, and the industry is growing at more than 30% per annum (*2*). However, low-cost, base-loadable, fossil-based electricity has always served as a formidable cost competitor for electrical power generation. To provide a truly widespread primary energy source, solar energy must be captured, converted, and stored in a cost-effective fashion. Even a solar electricity device that operated at near the theoretical limit of 70% efficiency would not provide the needed technology if it were expensive and if there were no cost-effective mechanism to store and dispatch the converted solar energy upon demand (*3*). Hence, a complete solar-based energy system will not only require cost reduction in existing PV manufacturing methods, but will also require science and technology breakthroughs to enable, in a convenient, scalably manufacturable form, the ultralow-cost capture, conversion, and storage of sunlight.

One key step is the capture and conversion of the energy contained in solar photons. Figure 1 shows the fully amortized cost of electricity as a function of the efficiency and cost of an installed PV module (*2*, *4*). Because the total energy provided by the Sun is fixed over the 30-year lifetime of a PV module, once the energy conversion efficiency of a PV module is established, the total amount of "product" electricity produced by the module at a representative mid-latitude location is known for the lifetime of the system. The theoretical efficiency limit for even an optimal single band gap solar conversion device is 31%, because photons having energies lower than the absorption threshold of the active PV material are not absorbed, whereas photons having energies much higher than the band gap rapidly release heat to the lattice of the solid and therefore ultimately contain only a useful internal energy equal to that of the band gap (*2*). Small test cells have demonstrated efficiencies of >20%, with the remaining losses almost entirely due to small reflection losses, grid shading losses, and other losses at the 5 to 10% level that any practical system will have to some extent. Shipped PV modules now have efficiencies of 14 to 20% in many cases. At such an efficiency, if the cost of a module is $300/m^2 (*2*), and if we take into account the accompanying fixed costs in the so-called "balance of systems" (such as the inverter, grid connection, etc., which add a factor of ~2 to the total installed system cost), then the sale price of grid-connected PV electricity must be $0.25 to $0.30 per kilowatt-hour (kWh) to recover the initial capital investment and cost of money over the lifetime of the PV installation (*2*, *4*). Currently, however, at utility-scale electrical power generation costs are much less, with current and new installations costing $0.03 to $0.05 per kWh (*1*). Hence, for solar electricity to be cost-competitive with fossil-based electricity at utility scale, improvements in efficiency are helpful, but manufacturing costs must be substantially reduced.

In current manufacturing schemes for Si-based solar cells, the cost of the processed and purified Si is only about 10% of the final cost of the PV module. Some of the Si is lost in cutting up boules into wafers, and other costs are incurred in polishing the wafers, making the diffused junction in the Si into a photovoltaic device, fabricating the conducting transparent glass, masking and making the electrical contacts, sealing the cells, connecting the cells together reliably into a module, and sealing the module for shipment. Hence, in such systems, the energy conversion efficiency is at a premium so as to better amortize these other fixed costs involved with making the final PV module.

Improvements in efficiency above the 31% theoretical limit are possible if the constraints that are incorporated into the so-called Shockley-Queisser theoretical efficiency limit are relaxed (*2*). For example, if photons having energies greater than the band gap of the absorbing material did not dissipate their excess energy as heat, but instead produced more voltage or

Beckman Institute and Kavli Nanoscience Institute, 210 Noyes Laboratory, 127-72, California Institute of Technology, Pasadena, CA 91125, USA. E-mail: nslewis@caltech.edu

generated multiple, low-energy, thermalized electrons from the energy of a single absorbed photon, theoretical efficiencies in excess of 60% would, in principle, be attainable. Absorbers having a highly quantized band structure, such as quantum wells and quantum dots, can theoretically produce the desired effects (Fig. 2). In fact, recent observations on PbSe quantum dots have demonstrated the production, with high quantum yield, of multiple excitons from a single absorbed photon, thereby establishing an existence proof for the process of interest (5). At present, however, there is no method for efficiently extracting the photogenerated carriers from the quantum dot structure to produce electricity in an external circuit. Materials with "mini-bands" or with "intermediate bands" also offer the possibility for ultrahigh energy conversion efficiency (2, 6). In this approach, different incident photon energies would promote absorption from different isolated energy levels and therefore allow for the production of different voltages (Fig. 2). The phenomenon has been described theoretically but has yet to be demonstrated in a practical implementation. In addition, these materials are currently extremely costly, and methods of retaining the high performance with scalable, inexpensive manufacturing methods would also be required.

In the absence of marked increases in cell efficiency, the value of new solar cell materials rests primarily with their potential to enable an entirely different manufacturing process, such as roll-to-roll manufacturing, printing, painting, or other ultralow-cost approaches to implementation of PV technology. This area is where breakthroughs in the science and technology of solar cell materials can have the greatest impact on the cost and widespread implementation of solar electricity.

The key issue involves the trade-off between material purity and device performance. In a typical planar solar cell design, the charge carriers are collected in the same direction as light is absorbed. A minimum thickness of the cell is set by the thickness of material required to absorb ~90% of the incident sunlight. However, the required thickness of the material also imposes a constraint on the required purity of the material, because the photoexcited charge carriers must live sufficiently long within the absorbing material to arrive at the electrical junction, where they can be separated to produce an electrical current flow through the metallic contacts to the cell. Impure absorber materials with short charge carrier lifetimes can therefore effectively absorb sunlight but cannot effectively convert that absorbed energy into electricity. In turn, absorber materials with the necessary purity are generally costly to produce and manufacture. Cheaper materials, such as organic polymers or inorganic particulate solids with small grain sizes, generally have short charge carrier lifetimes and/or induce recombination of charge carriers at the grain boundaries of such materials. This cost-thickness-purity constraint is largely why all current PV cells fall in the green region, labeled zone I, in Fig. 1.

Approaches to circumventing this cost-efficiency trade-off generally involve orthogonalization of the directions of light absorption and charge carrier collection. High-aspect ratio nanorods, for example, can provide a long dimension for light absorption while requiring only that carriers move radially, along the short dimension of the nanorod, to be separated by the metallurgical junction and collected as electricity (Fig. 3) (2, 7). A conceptually similar approach involves the use of interpenetrating networks of inorganic absorbers, such as CdTe "tetrapods" (8) and/or organic polymeric absorbers (9), such as the organic conducting polymer

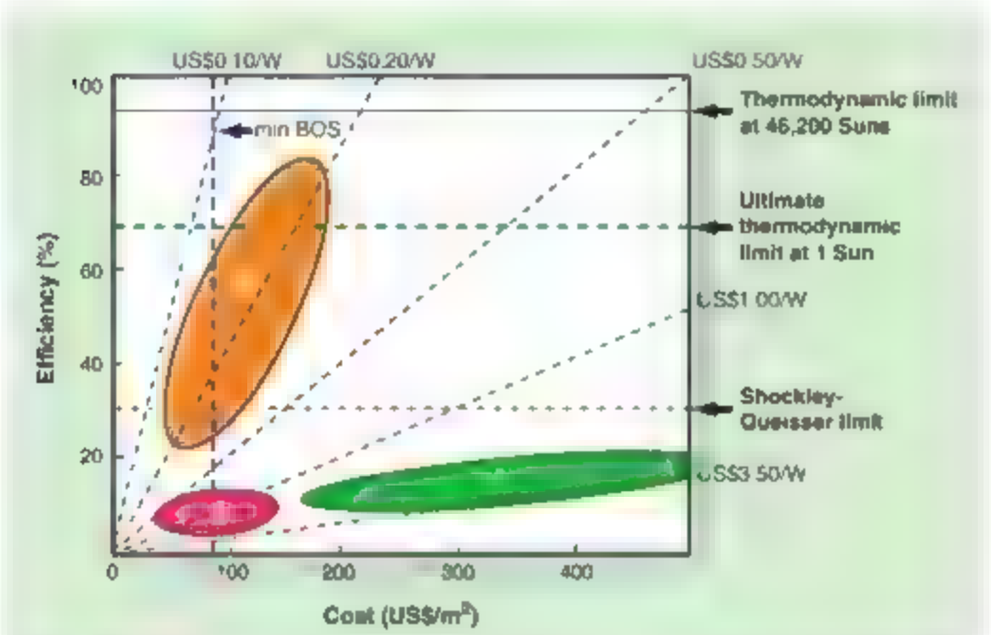


Fig. 1. Solar electricity costs as function of module efficiency and cost. The theoretical efficiencies are shown for three cases: the Shockley-Queisser limit for a quantum conversion device with a single band gap, in which carriers of lower energy are not absorbed and carriers of energy higher than the band gap thermalize to the band gap; the second-law thermodynamic limit on Earth for 1 Sun of concentration; and the second-law thermodynamic limit for any Earth-based solar conversion system. Current solar cell modules lie in zone I. The dashed lines are equal-cost lines on a cost per peak watt ($/$W_p$) basis. An estimate for the minimum balance-of-systems cost given current manufacturing methods is also indicated. A convenient conversion factor is that \$1/W_p amortizes out to ~\$0.05/kWh over a 30-year lifetime of the PV module in the field. [Adapted from (4)]

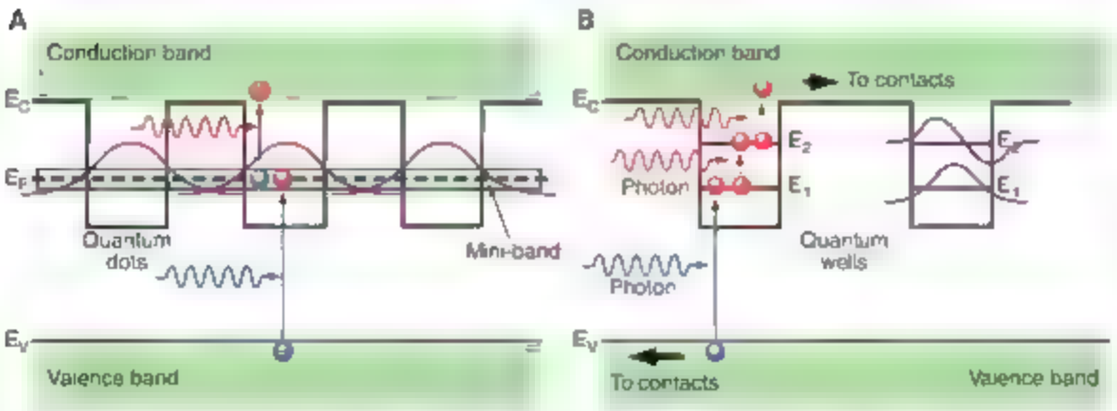


Fig. 2. Possible methods of circumventing the 31% efficiency limit for thermalized carriers in a single–band gap absorption threshold solar quantum conversion system. (**A**) Intermediate-band solar cell. (**B**) quantum-well solar cell. [Adapted from (2)]

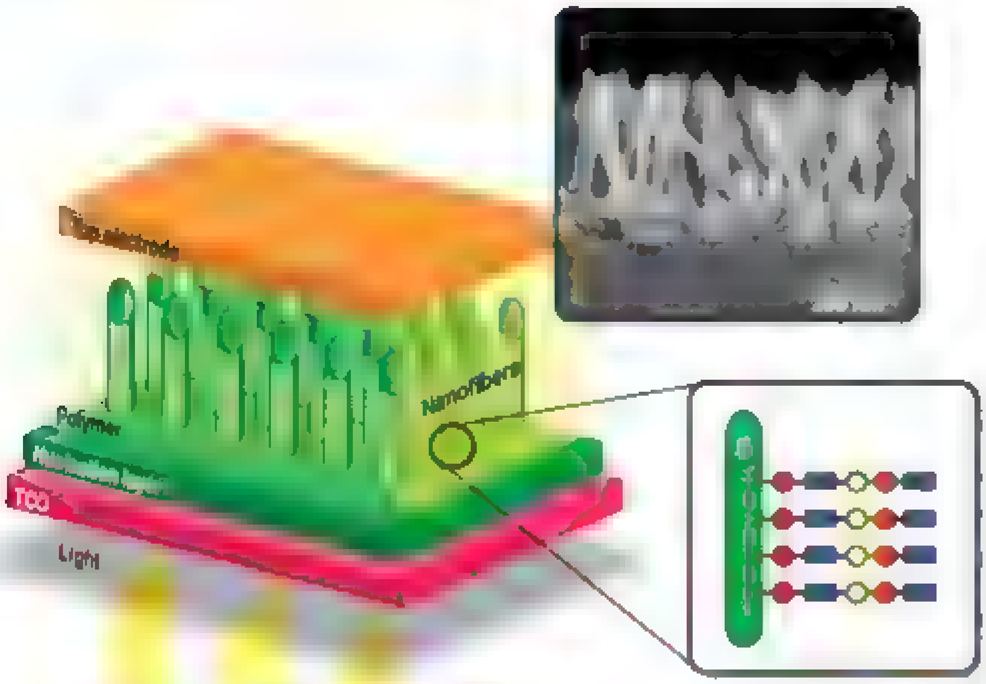


Fig. 3. Arrays of nanorods, illustrating an approach to orthogonalization of the directions of light absorption (down the length of the rods) and charge carrier collection (radially outward to the surface of the rods). [Adapted from (2)]

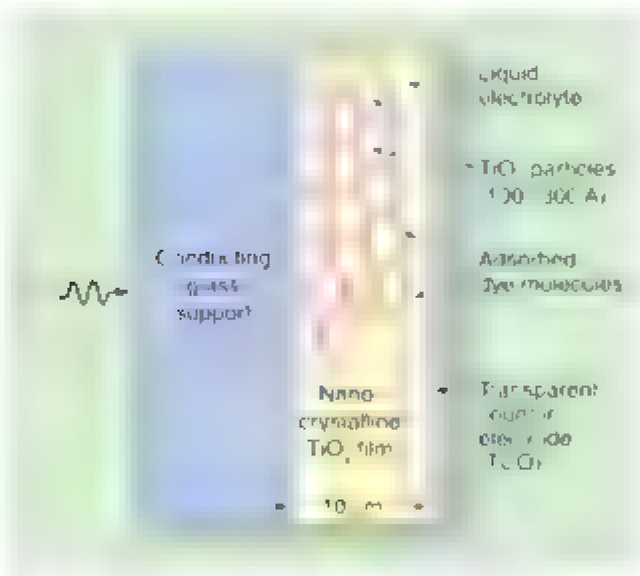


Fig. 4. Dye-sensitized solar cell, in which a nanoparticulate network provides collection of charge carriers injected into it as a result of absorption of sunlight by the adsorbed dye molecule. The oppositely charged carrier moves through the contacting liquid or polymeric phase to the counterelectrode, completing the electrical circuit in the solar cell. [Adapted from (2)]

poly(phenylenevinylene). Such systems are under widespread investigation at present, and the key is not only to obtain intimate contact between the light-absorbing and charge-conducting phases but also to control the chemistry at the interface between the two phases that make up the device. Junction recombination is a deleterious loss pathway even in many planar solar cell devices, and such junction recombination generally becomes dominant in disordered systems that, by definition, have a large increase in their interfacial contact area relative to their projected geometric area for light absorption. Methods for controlling the chemical properties of the surfaces and junctions of such systems, and thereby reducing their natural tendency to promote deleterious charge carrier recombination, are therefore critical. Such methods have been developed for certain well-defined semiconductor surfaces (10) and will need to be developed and implemented successfully for the high junction area systems to obtain high (>5%) energy conversion efficiencies from such devices.

A conceptually related system is the dye-sensitized solar cell, in which a random, disordered network of inexpensive TiO_2 particles is used to collect the charge carriers. The light absorption is performed by an adsorbed dye molecule, and the interfacial contact distance is kept small by use of a liquid or conductive polymer to penetrate the pore structure of the solid and collect the other charge carrier type to complete the circuit in the cell (Fig. 4) (11). Small "champion" dye-sensitized solar cells have shown efficiencies as high as 10 to 11%, although at present large-area devices typically have efficiencies of <5%. Improvements in the efficiency of such systems will require improved dyes, better electrolytes, and better control over the recombination at the interfacial contact area that currently limits the voltage produced by such systems to about 50 to 60% of its theoretical value. The stability of such systems will also need to be demonstrated under operational conditions for extended periods (>10 years) to allow them to be implemented in the marketplace. Clearly, advances in basic science are needed to enable all such nanostructured systems to truly offer a practical, ultralow-cost option for solar electricity production (2).

Although there is tremendous potential for growth for PV in electricity generation, solar electricity can never be a material contributor to primary energy generation without cost-effective methods for storing and distributing massive quantities of electricity (2, 3, 12). Put simply, the Sun goes out locally every night, and the intermittency imposed by the diurnal cycle must be dealt with to provide a full, base-loadable primary energy system from the Sun. The lack of cost-effective large-scale electrical storage capacity on Earth underlies the call for development of space-based solar power systems. On Earth, the cheapest method for massive electricity storage is pumped-water storage, which can be relatively efficient, but even that process does not scale well if every reservoir would have to be filled up each day and emptied each night. Additionally, a staggering amount of water would be needed to compensate for the diurnal cycle if one were to provide a material contribution to the primary U.S. or global energy generation through this approach. Batteries are a natural approach to electricity storage, but for battery storage to be cost-effective over the 30-year amortized lifetime of a PV system, enormous quantities of batteries would have to be hooked up to the grid, and they would have to cost as little as lead-acid batteries while providing the cycle life of lithium-ion batteries. Innovative approaches to massive, low-cost energy storage—including potentially a superconducting global transmission grid, supercapacitors, flywheels, etc., as promoted by Smalley (12)—will be important enablers of a full solar capture, conversion, and storage energy system.

Perhaps the most attractive method for cost-effective massive energy storage is in the form of chemical bonds (i.e., chemical fuel). After all, this approach is central to photosynthesis and is the basis for much of the recent attention devoted to development of biofuels. Photosynthesis, however, saturates at about one-tenth of the intensity of normal sunlight, and consequently the yearly averaged energy storage efficiency of even the fastest-growing plants is less than 1%, and typically less than 0.3 to 0.5% (2), as compared to the >15% efficiency values displayed by current PV devices (2). Hence, to first order, land-related constraints dominate the ultimate commercial potential of biofuels as

material contributors to primary energy supply whereas cost-related constraints dominate the ultimate commercial potential of PV-derived solar energy conversion and storage systems.

One approach to storing electrical energy in chemical bonds is through electrolysis, in which water is split into H_2 and O_2 in an electrolyzer. However, Pt-based electrolysis in acidic or neutral media is expensive and unlikely to be scalable to the levels that would be required for this process to be material in global primary energy production. Ni-based electrolysis in basic aqueous solutions is cheaper but requires scrubbing the input stream to remove the CO_2 (*13*); additionally, even the best fuel cells are only 50 to 60% energy efficient, and the best electrolysis units are 50 to 70% energy efficient (*13*), so the full-cycle energy storage/discharge efficiency of such a system is currently only 25 to 30%. Clearly, better catalysts for the multielectron transformations involved in fuel formation are needed. Nature provides the existence proof for such catalysts, with the hydrogenase enzymes operating at the thermodynamic potential for production of H_2 from H_2O, and with the oxygen-evolving complex of photosystem II producing O_2 from H_2O in an energy-efficient fashion. However, no human-made catalyst systems, either molecular or heterogeneous, have yet been identified that show performance even close to that of the natural enzymatic systems. Development of such catalysts would provide a key enabling technology for a full solar energy conversion and storage system.

Whether the fuel-forming system is separate as in a PV-electrolysis combination, or integrated, as in a fully artificial photosynthetic system that uses the incident charge-separated electron-hole pairs to directly produce fuels with no wires and with only water and sunlight as the inputs, is an interesting point of discussion from both cost and engineering perspectives. However, the key components needed to enable the whole system remain the same in either case: cost-effective and efficient capture, conversion, and storage of sunlight. Each of these functions has its own challenges, and integration of them into a fully functioning, synergistic, globally scalable system will require further advances in both basic science and engineering. Such advances, together with advances in existing technologies, will be required if the full potential of solar energy is to be realized.

References and Notes

1. *World Energy Assessment Overview, 2004 Update*, J. Goldemberg, T. B. Johansson, Eds. (United Nations Development Programme, New York, 2004) (www.undp.org/energy/weaover2004.htm).
2. *Basic Research Needs for Solar Energy Utilization* (U.S. Department of Energy, Washington, DC, 2005) (www.er.doe.gov/bes/reports/abstracts.html#SEU).
3. M. I. Hoffert *et al.*, *Science* **298**, 981 (2002).
4. M. A. Green, *Third Generation Photovoltaics: Advanced Solar Energy Conversion* (Springer-Verlag, Berlin, 2004).
5. R. D. Schaller, M. A. Petruska, V. I. Klimov, *Appl. Phys. Lett.* **87**, 253102 (2005).
6. M. A. Green, *Prog. Photovolt.* **9**, 137 (2001).
7. B. M. Kayes, H. A. Atwater, N. S. Lewis, *J. Appl. Phys.* **97**, 114302 (2005).
8. D. J. Milliron, I. Gur, A. P. Alivisatos, *MRS Bull.* **30**, 41 (2005).
9. G. Yu *et al.*, *Science* **270**, 1789 (1995).
10. N. S. Lewis, *Inorg. Chem.* **44**, 6900 (2005).
11. M. Grätzel, *Nature* **414**, 338 (2001).
12. R. Smalley, *Bull. Mater. Res. Soc.* **30**, 412 (2005).
13. J. Ivy, *Summary of Electrolytic Hydrogen Production* (National Renewable Energy Laboratory, Golden, CO, 2004) (www.nrel.gov/docs/fy04osti/36734.pdf).

14. Supported by the U.S. Department of Energy and NSF.

10.1126/science.1137014

PERSPECTIVES

Challenges in Engineering Microbes for Biofuels Production

Gregory Stephanopoulos

Economic and geopolitical factors (high oil prices, environmental concerns, and supply instability) have been prompting policy-makers to put added emphasis on renewable energy sources. For the scientific community, recent advances, embodied in new insights into basic biology and technology that can be applied to metabolic engineering, are generating considerable excitement. There is justified optimism that the full potential of biofuel production from cellulosic biomass will be obtainable in the next 10 to 15 years.

The idea of converting biomass-derived sugars to transportation biofuels was first proposed in the 1970s. Once again, the idea is being seriously contemplated as a possible substitute for petroleum-based liquid fuels. Economic and geopolitical factors (high oil prices, environmental concerns, and supply instability) have certainly played a role in reviving interest in renewable resources. However, an additional impetus is now provided by scientific and technological advances in biosciences and bioengineering that support increased optimism about realizing the full potential of biomass in the liquid fuels area within the next 10 to 15 years. New approaches to biology are being shaped by the genomics revolution, unprecedented ability to transfer genes, modulate gene expression, and engineer proteins, and a new mind-set for studying biological systems in a holistic manner (systems biology (*1*)). We are also seeing advances in metabolic engineering (*2*–*4*), with the goal of overproducing useful compounds by rationally and combinatorially engineering cells and their metabolic pathways (*5*). Combination of concepts and methods from these fields will create a platform of technologies that are critical for overcoming remaining obstacles in cost-efficient biofuel production from cellulosic biomass.

Figure 1 shows the basic features of a biomass-to-biofuels (B2B) process (*6*). After harvest, biomass is reduced in size and then treated to loosen up the lignin-cellulose fiber entanglement in a step that can take from a few minutes to many hours. Several methods have been used for this purpose, such as biomass treatment with saturated steam at 200°C, explosion with ammonia, and cooking with warm dilute acid (*6*). Dilute acid pretreatments are fast (minutes), whereas steam-based treatments can take up to a day. After pretreatment, the solid suspension is exposed to cellulolytic enzymes that digest the cellulosic and hemicellulosic biomass components to release the hydrolysis products, primarily six- and five-carbon sugars, respectively (along with acetic acid and lignin-derived phenolic by-products). The type of pretreatment defines the optimal enzyme mixture to be used and the composition of the hydrolysis products. The latter are fermented by ethanol-producing microorganisms such as genetically engineered yeasts, *Zymomonas mobilis* (Fig. 2), *Escherichia coli*, or *Pichia stipitis* (Fig. 3). Presently, cellulose hydrolysis and fermentation are combined in a single unit, termed the simultaneous saccharification fermentation (SSF) stage. The rationale of combining saccharification (the breaking up of complex carbohydrates into monosaccharides) and fermentation (the conversion of a carbohydrate to carbon dioxide and alcohol) in a single unit was to prevent inhibition of the hydrolytic enzymes by the reaction products (*7*). The SSF step typically lasts 3 to 6 days, with cellulose hydrolysis being the slow, limiting step. The product of SSF is a rather dilute ethanol stream of 4 to 4.5% from which ethanol is separated by distillation.

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: gregstep@mit.edu

Biomass pretreatment and hydrolysis are areas in need of drastic improvement. Despite substantial reduction in the cost of cellulolytic enzymes (*8*), sugar release from biomass still remains an expensive and slow step, perhaps the most critical in the overall process. Intensive research and development in all areas of enzyme production reduced the cost of cellulolytic enzymes by a factor of 50 to 80, down to 20 to 30 cents per gallon of ethanol produced (*8*, *9*). Although this is certainly an important advance, it is estimated that the enzyme cost will have to be further reduced to a level comparable to that of current approaches that produce ethanol from the starch in corn kernels at a cost of 3 to 4 cents per gallon of ethanol. Expression of cellulases in fermenting organisms or transfer of the biofuel-synthesizing pathway into a cellulase-producing organism are being pursued in a process termed the consolidated bioprocess (CBP) (*10*). CBP, however, is presently hampered by the relative inability of yeast to process recombinant cellulases at high rates through their endoplasmic reticulum and secretory pathways, and the relative (with regard to *E. coli* and yeast) lag in development of molecular biological methods to manipulate organisms such as *Trichoderma* that secrete cellulases naturally. The fact that glucose suppresses respiration in *Saccharomyces cerevisiae* reduces the amount of adenosine triphosphate available for protein biosynthesis, which may also render it difficult for enzyme production in yeast to be competitive with enzyme production by aerobic fungi such as *Trichoderma* or *Aspergillus*. When realized, CBP will enjoy the benefit of completely eliminating the cost of purifying cellulase and of higher activity of the cell-associated cellulase enzyme. To accomplish this goal, the hydrolysis and fermentation steps will have to be coordinated well inside a single cell, such that neither one limits the overall conversion process to proceed at maximum capacity. Although attainable over a longer time scale, in the near term B2B will benefit from the availability of large amounts of inexpensive and more active cellulases. This opportunity should be pursued by coordinated approaches from protein engineering, fungal overexpression, and bioprocess engineering to take advantage of economies of scale in enzyme production.

In engineering better microorganisms for biofuels production, combinatorial searches for promising target genes and other lab-scale experiments should be conducted with synthetic media. In terms of identified target genes or cell phenotypes, results obtained with the more convenient complex media (Luria broth or yeast extract) do not usually translate well to industrial conditions that use synthetic media.

Nonenzymatic, physicochemical hydrolysis methods (such as high-temperature pretreatments and hot-acid hydrolysis) are much faster than enzymatic approaches, albeit at the cost of reduced sugar yields due to undesirable side reactions. This is a problem that can be potentially solved by novel bioreactor designs operating at optimal contact times so as to maximize the rate of sugar-degrading side reactions without impairing biomass hydrolysis in the first place. The presence of lignin that effectively accumulates in the solids fraction as the carbohydrates are hydrolyzed away can interfere mechanically with filtration and recycling operations and complicate efforts to optimize the performance of the hydrolysis step. Advanced material-handling methods and new filtration devices specifically addressing the peculiarities of lignin consistency, or sequence-reverse schemes (whereby lignin removal precedes the hydrolysis step) are some possibilities that could exploit the fast rates of physicochemical hydrolysis while minimizing adverse side reactions. Finally, use of novel types of solvents such as those derived from ionic liquids are promising alternatives that should be further evaluated.

The cost competitiveness of a process such as that depicted in Fig. 1 depends on product titer, yield, and productivity. Final product titer is an important cost determinant not only because it affects the downstream purification cost but because it defines the size of the footprint of the entire processing plant. Low product titer is caused by various factors, including the total amount of substrate solids fed to the fermentor, the presence of inhibitory compounds as by-products of biomass hydrolysis (such as aromatics, furfurals, furan derivatives, and phenolics), and, of course, the toxicity of the final product itself. If, as seems likely, we can increase the solids loading into the SSF unit, then we may be able to increase substantially the final ethanol concentrations. This makes the engineering of ethanol-tolerant strains, which can tolerate the adverse environment in which the process takes place, of the utmost importance. Not much progress has been made on this front, perhaps because of the preconception that a complex phenotype such as ethanol tolerance could be modulated by a single gene, or at most a handful of genes. There is now accumulating evidence that no single gene can endow microbes with tolerance to ethanol and other toxic compounds. On the contrary, tolerance is a multigenic trait that must be elicited by drastically different approaches, such as global transcription machinery engineering (*11*). This method and its extensions should be systematically explored to identify transcription factor mutants that can increase the tolerance of industrial strains to the final fuel product as well as other relevant toxic compounds.

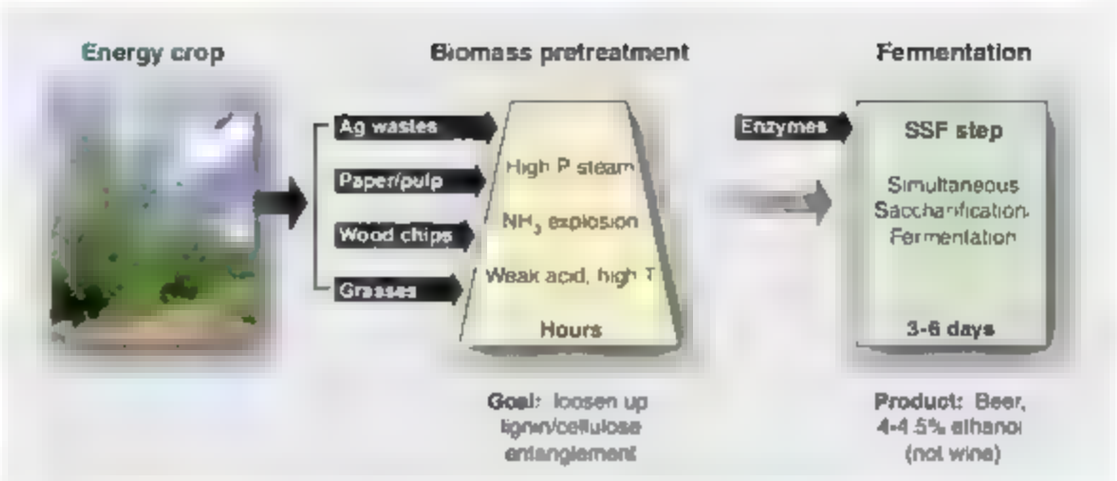


Fig. 1. Schematic of the overall conversion process of an energy crop to ethanol.

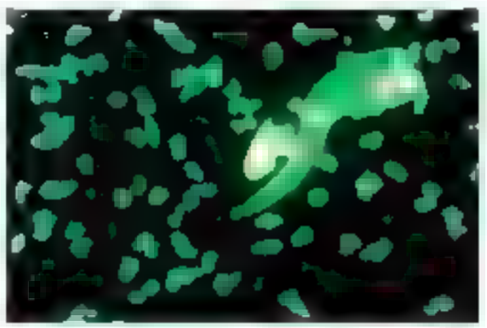


Fig. 2. *Zymomonas mobilis*, a metabolically engineered bacteria used for fermenting both glucose and xylose to ethanol. [Credit: Zhang Min; DOE/NREL]

Because the cost of a biomass-derived fuel depends critically on the yield of sugar conversion to the final product, much attention has been focused on the engineering of strains to use all sugars released from biomass hydrolysis, in particular the pentose sugars that are products of hemicellulose hydrolysis. Such sugars may constitute 5 to 30% of the total carbohydrates. Across various strategies have been used to attempt either to introduce the ethanol pathway in natural xylose consumers (*12*) or to engineer the

Fig. 3. An 8000-liter fermentation tank used to start the process of turning cellulosic material into ethanol. [Credit New Energy Company of Indiana; DOE/NREL]

xylose-catabolizing pathway in natural ethanol producers (*13*, *14*). The state of the art is rather well advanced as far as the engineering of various central ways (including pentose phosphate, glycolytic, ethanologenic, and redox balancing) is concerned. A variable that has received relatively limited attention is that of sugar transporters and their regulation. There is evidence that a multitude of such transporters may be in operation (*15*) and that their activity may depend on signaling defined by the sugar composition of the fermentation medium (*16*). Elucidation of sugar transport at the molecular level and better characterization of kinetic and regulatory properties, including quorum-sensing mechanisms, should be given high priority because they may provide the basis for the simultaneous use of the sugar mixtures released from biomass hydrolysis as opposed to the slower and suboptimal sequential use characterizing most present operations (sugars are consumed simultaneously in, for example, fermentations by recombinant *Z. mobilis* albeit at low rates). It is important to remember that one mole of CO_2 is produced for each mole of ethanol, for a total yield of 0.51 g of ethanol per gram of glucose consumed. As carbon oxidation to CO_2 is essential for generating the energy and redox equivalents needed to sustain cellular functions and the ethanol pathway itself, an interesting long-term idea is the capture and conversion to liquid fuels of this CO_2 by means of hydrogen supplied from carbon-free sources (such as nuclear or solar). This could be accomplished by conventional Fischer-Tropsch processes.

Process productivity is a principal determinant of capital cost. For cellulosic ethanol, the capital cost is estimated at ~$4 per gallon, contributing 30 to 35% of the ethanol manufacturing cost (*17*). However, these figures, as well as the others quoted in the following paragraph, vary considerably from source to source and are also time dependent. They should be viewed only as preliminary estimates that need to be validated by detailed empirical and analytic work. Furthermore, costs contributed by the process units are interrelated and cannot be assessed in isolation. Overall system analysis is critical for assessing the relative importance of the various process units and their interactions. Thus, reliable simulation packages for the integrated system operation must be developed for overall system analysis, optimization, and sensitivity studies (*18*).

The capital cost must be reduced by more than half for an economical process (along with a similar reduction in the feedstock cost, which will come primarily from yield improvements of an energy crop, and a 15 to 25 cents per gallon reduction in cellulolytic enzyme costs). Achieving the above goal or, equivalently, doubling process productivity, requires a coordinated approach for improving all units of the process and, in particular, the biomass pretreatment-hydrolysis steps mentioned earlier, because these are apparently the process rate-limiting steps. After that, the volumetric productivity of fermentation must be improved (presently between 1.5 and 2.0 grams of ethanol produced per hour and fermentor volume), which is the product of the specific productivity (grams of product produced per gram of cells per hour) and the total cell concentration that can be sustained in the fermentor. The latter again, is limited by the presence of the same inhibitory compounds; hence, use of more tolerant strains will affect the total process productivity as well. Additionally, specific productivity must be increased.

Various approaches have been suggested for increasing the specific ethanol productivity, such as increasing the amount of "rate-limiting" enzymes, enzyme deregulation, cofactor replenishment, and increase of precursor supply. Although some of these approaches are valid, some others are grossly misguided. A rather obvious approach to increase the flux through a pathway is by increasing the activity of every single enzyme in the pathway. Although this is acceptable for modest flux enhancements, it has not been attempted for large, order-of-magnitude scale flux increases on the grounds that it will cause large perturbations in the metabolites and, hence, the physiology of the organism. Yet, this will not happen if all enzymes in the pathway are similarly amplified, because the same steady state with respect to metabolite levels will be preserved. Simultaneous increase of the activity of all enzymes by a factor *f* will not affect metabolite levels, while allowing pathway flux to increase by the same factor *f*. The only limitation in such a scheme is the cell volume, which may not be able to accommodate drastically increased amounts of all enzymes of a pathway (it is estimated that the enzymes of glycolysis make up 10 to 15% of the total cellular protein). However, this problem can be overcome by engineering more active enzymes. Pathway flux amplification by coordinated activity enhancement of the rate-limiting enzymes (*19*) has been successfully used in lysine biosynthesis (*20*), aromatic amino acid production (*21*), and polyhydroxybutyrate synthesis in *E. coli* (*22*), among other systems. Determination of flux split ratios at key metabolic branch points (*23*), guided by flux determination methods (*24*, *25*), can aid this research, along with advanced fermentor feeding strategies that control metabolic activity (*26*).

Product separation for ethanol, the main biofuel currently produced, is carried out by distillation. Although mature and well optimized, it remains an energy-intensive and overall expensive step contributing 17 to 20 cents per gallon (*17*, *27*). In light of accumulating reports describing conformational changes in materials in response to small environmental changes (pH, temperature, or ionic strength), it may be useful to evaluate such phenomena with respect to their potential to facilitate ethanol separation. One can envision, for example, processes in which ethanol adsorbs preferentially on some material and desorbs when the material changes conformation after a small environmental change. Separation in such schemes would be entropically (as opposed to enthalpically) driven, could be less energy intensive than current operations, and possibly could be consolidated with fermentation in a single step.

As mentioned, ethanol is not the sole or optimal fuel to be produced from cellulosic biomass. Butanol is currently attracting attention because of its potential superior properties with respect to corrosiveness, volatility, energy density, and ease of separation (*28*). Aside from butanol, other higher alcohols, alkanes, and various types of oils are possible biochemically derived biofuels. It is not clear yet which one(s) will be the ideal biofuel, and the answer to this question may well depend on additional factors, such as the type of biomass available, particular climate conditions, and composition of engine emissions. The ability to clone, transfer, and control genes from different

organisms, including plants, has reached the point at which researchers will be able to engineer pathways that take advantage of a variety of conditions with a great degree of confidence. Additionally, genome sequences now provide a straightforward supply of genes to be tested in tentative pathway constructs. Nevertheless, it is important to develop technologies for the synthesis and separation of these alternative fuels, because it is yet unclear what additional requirements such technologies will pose in the design of a robust, cost-efficient, community-scale process.

In assessing the potential of current and projected technologies to develop cost-efficient B2B processes, it is important to bear in mind that the present state of affairs was reached by minimal investment directly in biofuels research. The major biosciences and bioengineering infrastructure was developed in the process of exploring medical applications of biology and biotechnology. Although this platform is the basis for the present optimism surrounding the use of biosciences for biofuel production from renewable resources, a number of problems still remain in realizing this potential. These problems must be addressed directly and adequately in the immediate future.

References and Notes

1. F. Ideker, T. Galitski, L. Hood, *Annu. Rev. Genomics Hum. Genet.* **2**, 343 (2001).
2. J. E. Bailey, *Science* **252**, 1668 (1991).
3. G. Stephanopoulos, J. J. Vallino, *Science* **252**, 1675 (1991).
4. G. Stephanopoulos, A. J. Sinskey, *Trends Biotechnol.* **11**, 392 (1993).
5. G. Stephanopoulos, *AIChE J.* **48**, 920 (2002).
6. B. Hahn-Hagerdal, M. Galbe, M. F. Gorwa-Grauslund, G. Liden, G. Zacchi, *Trends Biotechnol.* **24**, 549 (2006).
7. M. Takagi *et al.*, *Proceedings of the Bioconversion Symposium*, IIT Delhi, p. 551 (1977).
8. Department of Energy, *From Biomass to Biofuels: A Roadmap to the Energy Future*, based on a workshop (Rockville, MD, December 7 to 9, 2005).
9. Data from presentations of K. C. McFarland (Novozymes) and G. Ander (Genencor International) at the World Congress on Industrial Biotechnology and Bioprocessing, Orlando, FL, July 2006.
10. L. Lynd, *Curr. Opin. Biotechnol.* **16**, 577 (2005).
11. H. Alper, J. Moxley, E. Nevoigt, G. R. Fink, G. Stephanopoulos, *Science* **314**, 1565 (2006).
12. L. O. Ingram *et al.*, *Appl. Environ. Microbiol.* **53**, 2420 (1987).
13. M. Kuyper *et al.*, *FEMS Yeast Res.* **4**, 69 (2003).
14. M. Zhang *et al.*, *Science* **267**, 240 (1995).
15. E. Boles, C. P. Hollenberg, *FEMS Microbiol. Rev.* **21**, 85 (1997).
16. S. Ostan, M. Johnson, *Microbiol. Mol. Biol. Rev.* **63**, 554 (1999).
17. National Renewable Energy Laboratory, U.S. Department of Energy, "Softwood biomass to ethanol feasibility study," NREL Report SR-580-27310 (2004).
18. I wish to thank J. Deutch for valuable discussions on matters related to the cost of cellulosic ethanol and for pointing out the need to consider the costs of an ethanol process in its entirety and not of isolated units.
19. G. Stephanopoulos, T. W. Simpson, *Chem. Eng. Sci.* **52**, 2607 (1997).
20. M. Koffas *et al.*, *Metab. Eng.* **5**, 32 (2003).
21. R. Patnaik, R. G. Spitzer, J. C. Liao, *Biotechnol. Bioeng.* **46**, 361 (1995).
22. S. J. Sim *et al.*, *Nat. Biotechnol.* **15**, 63 (1997).
23. J. J. Vallino, G. Stephanopoulos, *Biotechnol. Prog.* **10**, 327 (1994).
24. M. I. Klapa *et al.*, *Biotechnol. Bioeng.* **62**, 375 (1999).
25. B. Folliard *et al.*, *Biotechnol. Bioeng.* **63**, 675 (1999).
26. R. D. Kiss, G. Stephanopoulos, *Biotechnol. Prog.* **7**, 501 (1991).
27. USDA, *2002 Ethanol Cost-of-Production Survey*, Agricultural Economic Report No. 841 (USDA, Washington, DC, 2005).
28. A BP-DuPont joint venture announced in June 2006 is focusing on biobutanol.

29. I thank H. Alper, J. Nielsen, D. Ramkrishna, and in particular J. McMillan, J. Deutch, and L. Lynd for reading the manuscript and providing many valuable comments.

10.1126/science.1139612

PERSPECTIVE

Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production

Michael E. Himmel,[1]* Shi-You Ding,[1] David K. Johnson,[1] William S. Adney,[1] Mark R. Nimlos,[1] John W. Brady,[2] Thomas D. Foust[3]

Lignocellulosic biomass has long been recognized as a potential sustainable source of mixed sugars for fermentation to biofuels and other biomaterials. Several technologies have been developed during the past 80 years that allow this conversion process to occur, and the clear objective now is to make this process cost-competitive in today's markets. Here, we consider the natural resistance of plant cell walls to microbial and enzymatic deconstruction, collectively known as "biomass recalcitrance." It is this property of plants that is largely responsible for the high cost of lignocellulose conversion. To achieve sustainable energy production, it will be necessary to overcome the chemical and structural properties that have evolved in biomass to prevent its disassembly.

High worldwide demand for energy, unstable and uncertain petroleum sources, and concern over global climate change have led to a resurgence in the development of alternative energy that can displace fossil transportation fuel. In response, many countries have initiated extensive research and development programs in biofuels, a sustainable and renewable energy resource that can provide liquid transportation fuels (*1*). The U. S. Department of Energy Office of the Biomass Program has developed a scenario for supplying 30% of the 2004 motor gasoline demand with biofuels by the year 2030, which roughly translates to a target of 60 billion gallons per year on a British thermal unit adjusted basis (*2*, *3*). Similarly, the European Union has developed a vision in which one-fourth of the E.U.'s transportation fuels will be derived from biofuels by 2030 (*4*). These political timetables result in critical challenges to the scientific community that require cutting-edge tools in the fields of systems and synthetic biology (*5*).

Starch from corn grain and simple sugars from sugar cane and beets are currently being used directly for ethanol fermentation, but to harness the structural sugars contained in plant fibers, we must first overcome the problems caused by biomass recalcitrance. Cellulose processing cannot commence until we improve (i) the relatively slow kinetics of breaking down pure cellulose into sugars, (ii) the low yields of sugars from other plant polysaccharides, and (iii) the removal of lignin, a relatively intractable polymer of phenylpropanoid subunits. It is clear that technological advances must be realized to make biofuels sustainable and cost effective.

In future biorefineries, biofuels will be produced from biomass resources, including corn grains and lignocellulosic biomass (such as agricultural residues, forestry wastes and thinnings, waste paper, and energy crops). Currently in the United States, approximately 455 million acres are in agricultural production to meet our food, feed, and fiber needs (*6*). A recent report (*7*) has suggested that in the near term, more than 1.3 billion tons of biomass could be produced annually in the United States on a sustainable basis, mostly from agricultural and forestry sources. Tilman and co-workers (*8*) have also described the potential role for low-input, high-diversity grassland perennials for bioconversion. Another study (*9*) has shown that biomass has the potential to simultaneously meet the nation's needs for liquid transportation fuel and for food, feed, and fiber, provided that we develop more advanced technologies and make certain land-use changes that would not require more net

[1]Chemical and Biosciences Center, National Renewable Energy Laboratory, Golden, CO 80401, USA. [2]Department of Food Science, Cornell University, Ithaca, NY 14853, USA. [3]National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO 80401, USA.

*To whom correspondence should be addressed. E-mail: mike.himmel@nrel.gov

land. The cost-competitive production of biofuels is currently prevented by the high cost of biomass feedstocks and the processes for converting biomass to sugars—that is, the cost of the thermochemical pretreatment and enzymatic hydrolysis unit operations in a biorefinery. Maximizing conversion yield is essential for offsetting feedstock cost.

Biomass Recalcitrance

Plant biomass has evolved complex structural and chemical mechanisms for resisting assault on its structural sugars from the microbial and animal kingdoms (Fig. 1). Natural factors believed to contribute to the recalcitrance of lignocellulosic feedstocks to chemicals or enzymes include (i) the epidermal tissue of the plant body, particularly the cuticle and epicuticular waxes; (ii) the arrangement and density of the vascular bundles; (iii) the relative amount of sclerenchymatous (thick wall) tissue; (iv) the degree of lignification (*10*); (v) the structural heterogeneity and complexity of cell-wall constituents such as matrix polysaccharides and matrix polymers (*11*); (vi) the cell linkages for enzymes acting on an insoluble substrate (*12*); and (vii) the inhibitors to subsequent fermentations that exist naturally in cell walls or are generated during conversion processes (*13*). In the context of the biorefinery, these chemical and structural features of biomass affect liquid penetration and/or enzyme accessibility and activity and, thus, conversion costs.

At the molecular level (Fig. 2), the crystalline cellulose core of cell-wall microfibrils (*14*) is highly resistant to chemical and biological hydrolysis because of its structure, in which chains of cellodextrins are precisely arranged. The chair conformation of the glucose residues in cellulose orders the hydroxyl groups into radial (equatorial) orientation and the aliphatic hydrogen atoms into axial positions. As a result, there is strong intersheet hydrogen bonding between adjacent chains in a cellulose sheet and weaker hydrophobic interactions between cellulose sheets. The hydrophobic face of cellulose sheets makes crystalline cellulose resistant to acid hydrolysis because it contributes to the formation of a dense layer of water near the hydrated cellulose surface (*15*). The strong intermolecular hydrogen-bonding network makes crystalline cellulose resistant to enzymatic hydrolysis (*14*), whereas hemicellulose and amorphous cellulose are readily digestible. Higher-order structures in plants also contribute to biomass recalcitrance. For example, access to the crystalline cellulose cores of microfibrils is restricted by a coating of amorphous cellulose and hemicellulose (*16*). At a microscopic and macroscopic scale, the complex heterogeneous nature of biomass creates mass transport limitations for delivery of chemical or biochemical catalysts.

Current Biomass Conversion Technology

The biorefinery is envisioned to comprise four major sections: feedstock harvest and storage, thermochemical pretreatment, enzymatic hydrolysis, and sugar fermentation to ethanol or other fuels. Existing biomass conversion schemes typically rely on a combination of chemical and enzymatic treatments. A pretreatment step is usually conducted to reduce recalcitrance by depolymerizing and solubilizing hemicellulose (approximately 20 to 40% weight by weight of biomass). This step converts hemicelluloses to monosaccharides and oligosaccharides, which can be further hydrolyzed or fermented. Removal of hemicellulose from the microfibrils is thought to expose the crystalline cellulose core, which can then be hydrolyzed by cellulase enzymes. In addition, pretreatment typically breaks down the macroscopic rigidity of biomass and decreases the physical barriers to mass transport.

Pretreatment. Thermochemical pretreatment of biomass has long been recognized as a critical technology to produce materials with acceptable enzymatic digestibilities. For example, dilute sulfuric acid pretreatment at 140° to 200°C renders the cellulose in cell walls more accessible to saccharifying enzymes. At moderate severities (*17*), the hemicelluloses are hydrolyzed and the sugars are solubilized as monomers and oligomers; however, the yields of solubilized sugars are less than quantitative (i.e., 60 to 70%) (*18*). For the acid treatments, release of monomeric and oligomeric sugars from hemicellulose exhibits multimodal kinetics in which a slow component directly relates to the high cost of conversion (*19*, *20*). For example, a number of researchers (*20*–*23*) have noted that the solubilization of xylan in hemicellulose appears to be best modeled as a pair of parallel first-order reactions: one that takes place at a fast rate and another that progresses at a much slower rate.

What governs this result is not clear at this time and it may depend on a number of factors, such as hemicellulose composition; biomass density; the presence of nonsugar components (such as lignin, acid neutralizing ash, and acetyl and other carboxylic acid groups); plant cell structure (including the types of cells or ratios of primary and secondary cell walls); or mass transport. Pretreatment schemes based on alkaline explosive decompression and organic solvent extractions have been proposed with considerable success (*13*). The alkaline process, known as ammonia fiber expansion (AFEX), leaves the hemicellulose in place but renders the remaining cell walls considerably more amenable to enzyme hydrolysis (*26*).

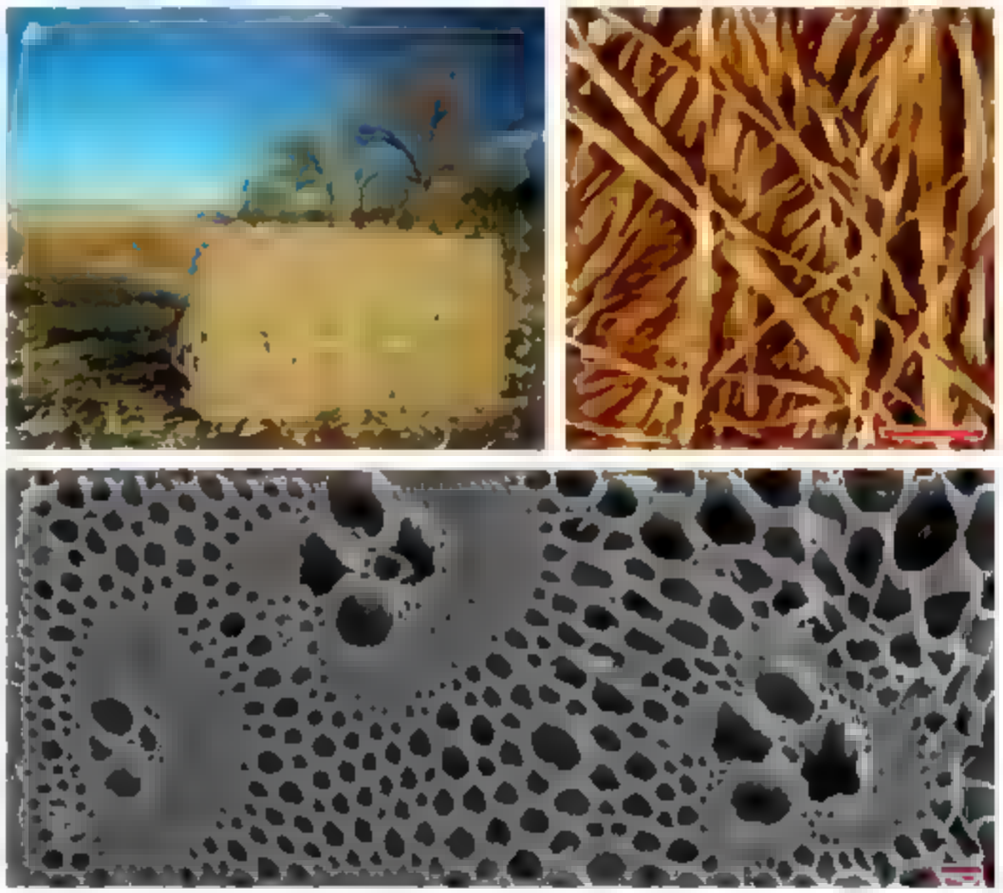


Fig. 1. Structural and chemical complexity of cell wall biomass. (**A**) Example of high-density bales of corn stover harvested on the eastern plains of northern Colorado. (**B**) An atomic force micrograph of the maize parenchyma cell wall surface. The diameter of individual microfibrils is only about 3 to 5 nm. Scale bar, 50 nm. (**C**) A scanning electron micrograph of the cross-section of a maize stem shows vascular bundles and pith tissues as well as the diverse cell sizes, shapes, and cell-wall thicknesses typical for higher-plant structure. Scale bar, 50 μm.

Enzymatic degradation. In nature, various cellulolytic microorganisms produce enzymes that function synergistically and associate with the microorganism (such as the cellulosome (*27*, *28*)) or act independently (such as most fungal and many bacterial cellulases) (*29*). Although it is not fully known how many enzymes are involved in cell-wall deconstruction, three general categories of enzymes are considered necessary to hydrolyze native cell-wall materials: cellulases, hemicellulases, and the accessory enzymes, which include hemicellulose debranching, phenolic acid esterase, and possibly lignin degrading and modifying enzymes (*29*). Once the hemicellulose barrier associated with cell-wall microfibrils has been compromised by chemical pretreatments, cellulase enzymes can be used to hydrolyze the crystalline cellulose cores of these structures.

Crystalline cellulose is hydrolyzed by the synergistic action of endo-acting (with respect to the cellulose chain) enzymes known as endoglucanases, and exo-acting enzymes, known as exoglucanases. The endoglucanases locate surface sites of activation, probably found at random, along the cellodextrin and insert a water molecule in the β (1,4) bond, creating a new reducing and nonreducing chain end pair. β-D-glucosidases (cellobiases) act to hydrolyze cellobiose, the product of cellulase action, and thus relieve the system from end-product inhibition. Cellulases and other glycosyl hydrolases (*30*) are known to proceed through a two-step, Koshland-type mechanism that leaves the terminal C1 carbon hydroxyl in the β configuration (retention of stereochemistry) or a concerted reaction mechanism that leaves the terminal hydroxyl in the α configuration (inversion of stereochemistry) (*31*). Water molecules could invade the space under the nonreducing chain end and thus prevent it from reannealing into the cellulose crystal. The removal of cellodextrins from the microfibril core is thought to occur at these new chain ends and this process, considered to be the rate-limiting step in cellulase action, is accomplished by exoglucanases also known as the "processive" cellulases.

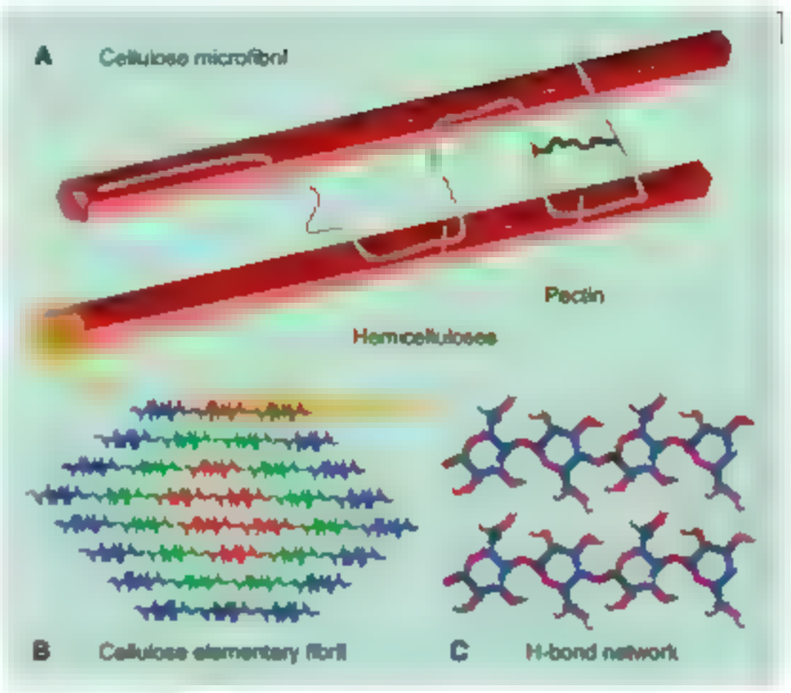


Fig. 2. (**A**) A simplified model showing the interaction of the major polysaccharides in the cell wall. (Lignin is not shown here because its interactions are not well established.) In this system, hemicelluloses are closely associated to the surface of the rigid cellulose crystallite forming the microfibril network. Pectins are cross-linked polysaccharides forming a hydrated gel that "glues" the cell-wall components together. (**B**) The 36-chain model of the cellulose elementary fibril. Here, the depiction of the glucan chains is based generally on an x-ray structure of cellulose Iβ (*39*). It has been proposed (*16*) that the cellulose elementary fibril may contain three groups of glucan chains: in group C1 (red) are six true crystalline chains; in group C2 (green) are 12 subcrystalline chains with a small degree of disorder; and in group C3 (blue) are 18 surface chains that are subcrystalline with a large degree of disorder. (**C**) The intra- and interchain hydrogen-bond network in cellulose Iβ.

Fig. 3. Artistic concept of an exoglucanase (the *T. reesei* cellobiohydrolase I) acting on crystalline cellulose. In this depiction, the carbohydrate-binding module (left) recognizes and binds to the cellulose surface. By a process not fully understood, a single chain of cellulose is "decrystallized" and directed into the active-site tunnel of the catalytic domain (right). This enzyme is thought to proceed along a cellulose chain cleaving one cellobiose unit per catalytic event until the chain ends or the enzyme becomes inactivated (*40*, *41*).

Overcoming Biomass Recalcitrance

Current biomass-conversion technologies are primarily developed empirically based on limited understanding of the biological and chemical properties of biomass. Recent studies of plant development, carbohydrate chemistry, and the ultrastructure of cell walls continue to provide new insights into biomass conversion. To reach the goal of producing cost-competitive biofuels from biomass, these new findings from plant science and carbohydrate chemistry must be translated and integrated into the conversion processes. Further studies will undoubtedly rely on, for example, the development of new techniques for imaging and characterizing the chemical topography of the cell wall at the nanometer scale. The nature of research aimed at overcoming biomass recalcitrance will primarily focus on the coengineering of new cell walls to be degraded by newly engineered enzymes designed for this role.

Plants designed for deconstruction. Recent studies of plant cell-wall biosynthesis are beginning to provide new understanding about the structure and chemistry of the plant cell wall (*10*). Although much of our knowledge is anecdotal, the cell walls of higher plants are viewed as an assembly of biopolymers, in some ways mimicking a "liquid crystal" synthesized by pathways with as-yet undetermined controls (*10*, *32*). For example, cellulose is synthesized and assembled on plasma membrane, whereas hemicelluloses are synthesized in the Golgi apparatus.

Despite our lack of detailed knowledge regarding cell-wall structure research during the past 20 years, largely reductionist in approach, has led to a body of information regarding treatments of the cell wall that are effective for enhancing enzyme action. Studies have shown that systematic removal of hemicelluloses, by either acidic or enzymatic processes, results in the marked reduction in cellulase loadings required to convert cellulose

to cellobiose or glucose (*33*, *34*). Other studies have shown that a reduction in phenolic esters, such as those characteristic of the linkages between lignins and hemicellulose, also permit a reduction in cellulase loadings (*11*). It has been theorized, therefore, that engineering plant cell walls by altering the molecular interactions between hemicelluloses, lignin, and cellulose microfibrils could result in more efficient use of costly cellulase enzymes. Finally, of considerably greater potential benefit, and accordingly greater difficulty, is the possibility of changing the nature of cellulose itself. Could the cellulose synthase complex be altered to produce "wounded" (in terms of either degree of crystallization or polymerization) cellulose more amenable to deconstruction? Would such a plant survive and thrive?

Another scale of effective engineering of plant cells and structure may be at the level of the organization of plant tissues. For example, because of the high economic and energetic costs of grinding, biomass-processing biorefineries would ideally be capable of using feedstocks in the 1- to 2.5-centimeter fragment range (*18*). This requires that catalyst (chemical and enzyme) penetration throughout the material be optimal. Diffusion of dissolved solids and water throughout plant tissue is controlled by the arrangement of vascular bundles, as well as pits between connecting cell walls. Therefore, another approach, one aimed at enhancing the penetration of pretreatment chemicals and hydrolytic enzymes, could be genetic engineering of the organization of vascular bundles and cell-wall pit density. Again, the concerns regarding plant survivability and vigor can be raised with such an approach.

Engineering catalysis and bioconversion systems. In many ways, processive cellulases are "protein machines" (Fig. 3); however, cellulase enzymes function about one to two orders of magnitude more slowly than other polysaccharidases (*35*). Unfortunately, the enzymatic decrystallization process is both critical and poorly understood, and thus considerable research is needed to enhance the performance of cellulase action. Efforts to improve the performance of cellulases primarily follow two courses, (i) mining diversity to find new enzyme paradigms, and (ii) knowledge-based protein engineering. For the latter approach to be effective, the mechanism of action of these enzymes in the context of the cellulose surface must be understood at the molecular level. Although attempts have been made to adapt directed evolution solutions to the problem of cellulase-specific activity improvement, limitations in effective high-throughput strategies and appropriate expression systems have slowed progress (*12*).

Future process scenarios have been proposed that combine key process steps, thus reducing overall process complexity and cost. One notable example is the consolidated biomass processing (CBP) technology (*36*). The CBP concept was probably initiated with the advent of the simultaneous saccharification and fermentation (SSF) scheme used by Gauss in a process developed for Gulf Oil (*37*). More recently, thoughts about combining SSF with enzyme production have resulted in new approaches to CBP, which could either require engineering an ethanologen (such as *Saccharomyces cerevisiae*) to be cellulolytic or engineering a cellulase producer (such as *Clostridium thermocellum*) to be ethanologenic. For the *C. thermocellum* case, the bioenergetic benefits specific to growth on cellulose result from the efficiency of oligosaccharide uptake combined with intracellular phosphorolytic cleavage of β-glucosidic bonds, another pathway not known in fungi. Scientists believe that these benefits exceed the bioenergetic cost of cellulase synthesis, supporting the feasibility of anaerobic processing of cellulosic biomass without added saccharolytic enzymes (*38*).

Outlook for an Advanced Biorefinery Industry

Ultimately, biomass conversion processes are attractive because they are in practice today and extension to future scenarios is easy for the public to envision. Although developing the technology for cost-effective motor fuel production by 2030 is challenging, the advances in scientific understanding necessary to achieve this goal appear realizable. The general path forward along the biological fuels production route will generally rely on consolidation of processing steps, both in the engineering and biological sense. Microbial cells will be expected to conduct multiple conversion reactions with high efficiency and to remain robust to process conditions. These improvements require deeper understanding of cellular and metabolic processes. New generations of hydrolytic enzymes will function near their theoretical limits, and energy plants will be modified to serve as improved substrates for these new generation enzymes. Indeed, it is entirely possible that the next generation of energy plants will harbor the genes encoding enzymes necessary for self-deconstruction, activated before harvest or at the normal conclusion of the growth cycle.

References and Notes

1. A. P. C. Faaij, *Energy Policy* **34**, 322 (2006).
2. U.S. Department of Energy, Office of the Biomass Program, 30 x 30 Workshop, Washington, DC, 1 to 2 August, 2006 (www.30x30workshop.biomass.govtools.us/).
3. Office of the Biomass Program (OBP), "Multi-Year Program Plan, 2007-2012" (OBP, U.S. Department of Energy, Washington, DC, 2005) (http://www1.eere.energy.gov/biomass/pdfs/mypp.pdf).
4. Biofuels Research Advisory Council, "Biofuels in the European Union: A Vision for 2030 and Beyond" (2006); available online (www.biomatnet.org/publications/191rep.pdf).
5. J. Houghton, S. Weatherwax, J. Ferrell, "Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda," Biomass to Biofuels, Rockville, MD, 7 to 9 December 2005 (to order the report, see http://genomicsgtl.energy.gov/biofuels/b2bworkshop.shtml#orderform).
6. U.S. Department of Agriculture–National Resource Conservation Service, National Resources Inventory: 2001 Annual NRI (NRCS, USDA, Washington, DC, 2003) (www.nrcs.usda.gov/technical/land/nri01/nri01lu.html).
7. R. D. Perlack *et al.*, "Biomass as Feedstock for a Bioenergy and Bioproducts Industry: The Technical Feasibility of a Billion-Ton Annual Supply" (Oak Ridge National Laboratory Report TM-2005, under contract DOE/GO-102005-2135, Oak Ridge, TN, (2005).
8. D. Tilman, J. Hill, C. Lehman, *Science* **314**, 1598 (2006).
9. S. Monhebel, *Renew. Sustain. Energy Rev.* **9**, 191 (2005).
10. D. J. Cosgrove, *Nat. Rev. Mol. Cell Biol.* **6**, 850 (2005).
11. K. Iyama, T. B. T. Lam, B. A. Stone, *Plant Physiol.* **104**, 315 (1994).
12. M. E. Himmel, M. F. Ruth, C. E. Wyman, *Curr. Opin. Biotechnol.* **10**, 358 (1999).
13. C. E. Wyman *et al.*, *Bioresour. Technol.* **96**, 1959 (2005).
14. Y. Nishiyama, P. Langan, H. Chanzy, *J. Am. Chem. Soc.* **124**, 9074 (2002).
15. J. F. Matthews *et al.*, *Carbohydr. Res.* **341**, 138 (2006).
16. S. Y. Ding, M. E. Himmel, *J. Agric. Food Chem.* **54**, 597 (2006).
17. Pretreatment severity is defined as the combined effect of temperature, acidity, and duration of treatment. Hot water pretreatments, which actually use acetic acid liberated from cell-wall hemicellulose, represent the lowest degree of severity.
18. R. Torget, M. E. Himmel, K. Grohmann, *Bioresour. Technol.* **35**, 239 (1991).
19. C. G. Liu, C. E. Wyman, *Ind. Eng. Chem. Res.* **43**, 2781 (2004).
20. A. Esteghlalian, A. G. Hashimoto, J. J. Fenske, M. H. Penner, *Bioresour. Technol.* **59**, 129 (1997).
21. M. Shiang, J. C. Linden, A. Mohagheghi, K. Grohmann, M. E. Himmel, *Biotechnol. Prog.* **7**, 315 (1991).
22. S. B. Kim, Y. Y. Lee, *Biotechnol. Bioeng. Symp.* **17**, 71 (1987).
23. M. T. Maloney, T. W. Chapman, A. J. Baker, *Biotechnol. Bioeng.* **27**, 355 (1985).
24. O. Mayans *et al.*, *Structure* **5**, 677 (1997).
25. Y. Sun *et al.*, *Bioresour. Technol.*, in press.
26. F. Peymoun, L. Laureano-Perez, H. Alizadeh, B. E. Dale, *Bioresour. Technol.* **96**, 2014 (2005).
27. E. A. Bayer, J. P. Belaich, Y. Shoham, R. Lamed, *Annu. Rev. Microbiol.* **58**, 521 (2004).
28. E. A. Bayer, L. J. W. Shimon, Y. Shoham, R. Lamed, *J. Struct. Biol.* **124**, 221 (1998).
29. M. E. Himmel *et al.*, in *Fuels and Chemicals from Biomass*, B. C. Saha, J. Woodward, Eds., vol. 666 of *American Chemical Society Symposium Series* (ACS, Washington, DC, 1997), pp. 2–45.
30. CAZy (Carbohydrate Active Enzymes) is a database for the families of structurally related catalytic and carbohydrate-binding modules of enzymes that degrade, modify, or create glycosidic bonds. The database is available online (www.cazy.org/CAZY/).
31. D. E. Koshland, *Biol. Rev.* **28**, 416 (1953).
32. M. S. Doblin, I. Kurek, D. Jacob-Wilk, D. P. Delmer, *Plant Cell Physiol.* **43**, 1407 (2002).
33. R. Torget, M. Himmel, K. Grohmann, *Appl. Biochem. Biotechnol.* **34/35**, 115 (1992).
34. B. Yang, C. E. Wyman, *Biotechnol. Bioeng.* **86**, 88 (2004).
35. K. Kipper, P. Valjamae, G. Johansson, *Biochem. J.* **385**, 527 (2005).
36. L. R. Lynd, P. J. Weimer, W. H. van Zyl, I. S. Pretorius, *Microbiol. Mol. Biol. Rev.* **66**, 506 (2002).
37. W. Gauss, U.S. Patent 3, 990, 944 (1976).
38. Y. Lu, Y. H. P. Zhang, L. R. Lynd, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16165 (2006).
39. Y. Nishiyama, J. Sugiyama, H. Chanzy, P. Langan, *J. Am. Chem. Soc.* **125**, 14300 (2004).
40. J. Rouvinen, T. Bergfors, T. Teeri, J. K. C. Knowles, T. A. Jones, *Science* **249**, 380 (1990).
41. M. Vrsanska, P. Biely, *Carbohydr. Res.* **227**, 19 (1992).

42. We acknowledge the support of the U.S. Department of Energy Office of the Biomass Program. We also thank W. Gretz for providing the picture shown in Fig. 1A, T. Vinzant for providing the image shown in Fig. 1B, and D. Seely at Pixel Kitchen for providing the image shown in Fig. 3.

10.1126/science.1137016

PERSPECTIVE

Ethanol for a Sustainable Energy Future

José Goldemberg*

Renewable energy is one of the most efficient ways to achieve sustainable development. Increasing its share in the world matrix will help prolong the existence of fossil fuel reserves, address the threats posed by climate change, and enable better security of the energy supply on a global scale. Most of the "new renewable energy sources" are still undergoing large-scale commercial development, but some technologies are already well established. These include Brazilian sugarcane ethanol, which, after 30 years of production, is a global energy commodity that is fully competitive with motor gasoline and appropriate for replication in many countries.

A sustainable energy future depends on an increased share of renewable energy, especially in developing countries. One of the best ways to achieve such a goal is by replicating the large Brazilian program of sugarcane ethanol, started in the 1970s.

The World Commission on Environment and Development (WCED) in 1987 defined "sustainable development" as development that "meets the needs of the present without compromising the ability of future generations to meet their own needs" (*1*). The elusiveness of such a definition has led to unending discussions among social scientists regarding the meaning of "future generations."

However, in the case of energy, exhaustible fossil fuels represent ~80% of the total world energy supply. At constant production and consumption, the presently known reserves of oil will last around 41 years, natural gas 64 years, and coal 155 years (*2*). Although very simplified, such an analysis illustrates why fossil fuels cannot be considered as the world's main source of energy for more than one or two generations. Besides the issue of depletion, fossil fuel use presents serious environmental problems, particularly global warming. Also, their production costs will increase as reserves approach exhaustion and as more expensive technologies are used to explore and extract less attractive resources. Finally, there are increasing concerns for the security of the oil supply, originating mainly from politically unstable regions of the world.

Except for nuclear energy, the most likely alternatives to fossil fuels are renewable sources such as hydroelectric, biomass, wind, solar, geothermal, and marine tidal. Figure 1 shows the present world energy use.

Fossil fuels (oil, coal, and gas) represent 80.1% of the total world energy supply, nuclear energy 6.3%, and renewables 13.6%. The largest part is traditional biomass (8.5% of total primary energy), which is used mainly in inefficient ways, such as in highly pollutant primitive cooking stoves used by poor rural populations, leading in many cases to deforestation.

The "new renewable energy sources" amount to 16 exajoules (1 EJ = 10^{18} J), or 3.4% of the total. Table 1 shows a breakdown of the contribution of new renewables, which include small hydropower plants. Many of these technologies are still undergoing large-scale commercial development, including solar, wind, geothermal, and modern biomass. The largest part (1.7% of the total) is modern biomass, which refers to biomass produced in a sustainable way and used for electricity generation, heat production, and transportation of liquid fuels. It includes wood and forest residues from reforestation and of sustainable management, as well as rural (animal and agricultural) and urban residues (including solid waste and liquid effluents).

From the perspective of sustainable energy development, renewables are widely available, ensuring greater security of the energy supply and reducing dependence on oil imports from politically unstable regions. Renewables are less polluting, both in terms of local emissions (such as particulates, sulfur, and lead) and greenhouse gases (carbon dioxide and methane) that cause global warming. They are also more labor-intensive, requiring more workforce per unit of energy than conventional fossil fuels (*3*).

Although technologically mature, some of the renewable sources of energy are more expensive than energy produced from fossil fuels. This is particularly the case for the "new renewables." Traditional biomass is frequently not the object of commercial transactions and it is difficult to evaluate its costs, except the environmental ones. Cost continues to be the fundamental barrier to widespread adoption of traditional biomass despite its attractiveness from a sustainability perspective.

A number of strategies have been adopted by governments in the industrialized countries and international financial institutions to encourage the use of "new renewables," and there have been several successes, based on the use of tax breaks, subsidies, and renewable portfolio standards (RPS). Examples are the large growth (of more than 35% per year, "albeit" from a low base value) for wind and solar photovoltaics in industrialized countries such as Denmark, Germany, Spain, and the United States (*4*). These technologies are slowly spreading to developing countries through several strategies.

In developing countries, the best example of a large growth in the use of renewables is given by the sugarcane ethanol program in Brazil. Today, ethanol production from sugarcane in the country is 16 billion liters (4.2 billion gallons) per year, requiring around 3 million hectares of land. The competition for land use between food and fuel has not been substantial. Sugarcane covers 10% of total cultivated land and 1% of total land available for agriculture in the country. Total sugarcane crop area (for sugar and ethanol) is 5.6 million hectares.

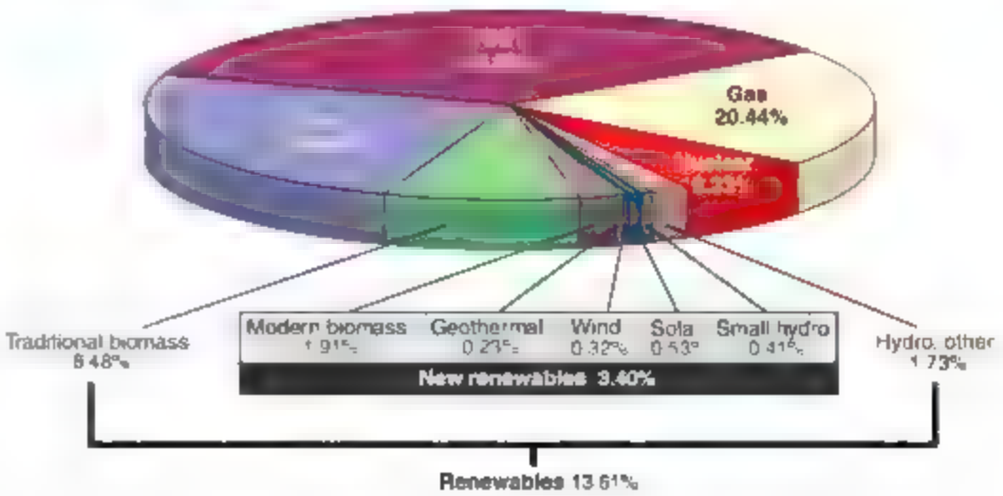


Fig. 1. World total primary energy supply 2004, shares of 11.2 billion tons of oil equivalent, or 470 EJ (*15, 16*).

University of São Paulo, São Paulo, Brazil. E-mail: goldemb@iee.usp.br

*Presently Secretary for the Environment, State of São Paulo, Brazil.

Production of ethanol from sugarcane can be replicated in other countries without serious damage to natural ecosystems. Worldwide, some 20 million hectares are used for growing sugarcane, mostly for sugar production (5). A simple calculation shows that expanding the Brazilian ethanol program by a factor of 10 (i.e., an additional 30 million hectares of sugarcane in Brazil and in other countries) would supply enough ethanol to replace 10% of the gasoline used in the world. This land area is a small fraction of the more than 1 billion hectares of primary crops already harvested on the planet.

What was the process that established firmly the ethanol program in Brazil? In the late 1970s, the Brazilian Federal Government mandated the mixture of anhydrous ethanol in gasoline (blends up to 25%) and encouraged car makers to produce engines running on pure hydrated ethanol (100%). Brazilian adoption of mandatory regulations determining the amount of ethanol to be mixed with gasoline (basically a Renewable Portfolio Standard for fuel) was essential to the success of the program. The motivation was to reduce oil imports that were consuming one-half of the total amount of hard currency from exports. Although it was a decision made by the federal government during a military regime, it was well accepted by the civil society, agricultural sector, and car manufacturers. Similar policies are being considered by the European Union, Japan, and several states in the United States.

Such a policy decision created a market for ethanol, and production increased rapidly. Ethanol costs declined along a "learning curve" (6) as production increased an average 6% per year from 0.9 billion gallons in 1980 to 3.0 billion gallons in 1990 and to 4.2 billion gallons in 2006. The cost of ethanol in 1980 was approximately three times the cost of gasoline, but governmental cross-subsidies paid for the price difference at the pump. The subsidies came mostly from taxes on gasoline and were thus paid by automobile drivers. All fuel prices were controlled by the government. Overall subsidies to ethanol are estimated to be around US$30 billion over 20 years (7), but were more than offset by a US$50 billion reduction of petroleum imports as of the end of 2006. Since the 1990s subsidies have been progressively removed, and by 2004 ethanol became fully competitive with gasoline on the international markets without government intervention. Subsidies for ethanol production are a thing of the past in Brazil (Fig. 2), because new ethanol plants benefit from the economies of scale and the modern technology available today, such as the use of high-pressure boilers that allow cogeneration of electricity, with surpluses sold to the electric power grid.

The Brazilian ethanol program started as a way to reduce the reliance on oil imports, but it was soon realized that it had important environmental and social benefits (8). Conversion to ethanol allowed the phasing-out of lead additives and MTBE (methyl tertiary butyl ether) and reduced sulfur, particulate matter, and carbon monoxide emissions. It helped mitigate greenhouse gas emissions efficiently, by having a net positive energy balance (renewable energy output versus fossil fuel inputs); also, sugarcane ethanol in Brazil costs less than other present technologies for ethanol production (Table 2) and is competitive with gasoline in the United States, even considering the import duty of US$0.54 per gallon and energy-efficiency penal-

Table 1. "New renewables," by source in 2004 (*15*); updated with data from (*4*, *16*). Assumed average conversion efficiency for biomass heat, 85%; biomass electricity, 22%; biomass combined heat and power (CHP), 80%; geothermal electricity 10%; all others, 100%.

Source/ technology	2004	
	Exajoules (EJ)	Share in this sector
Modern biomass energy		
Total	9.01	56.19%
Bioethanol	0.67	
Biodiesel	0.07	
Electricity	1.33	
Heat	6.94	
Geothermal energy		
Total	1.09	6.77%
Electricity	0.28	
Heat	0.30	
Small hydropower		
Total	1.92	12.00%
Wind electricity		
Total	1.50	9.35%
Solar		
Total	2.50	15.63%
Hot water	2.37	
Photovoltaic electricity, grid	0.06	
Photovoltaic electricity, off-grid	0.06	
Thermal electricity	0.01	
Marine energy (tidal)		
Total	0.01	
Total	16.03	100.00%

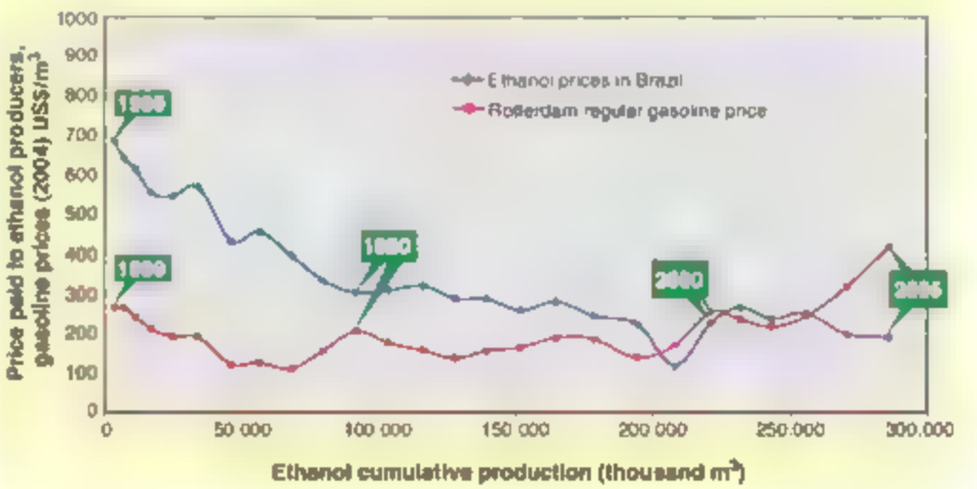


Fig. 2. Ethanol learning curve in volume, comparing the price paid to ethanol producers in Brazil with the price of gasoline in the international market of Rotterdam (*6*).

Table 2. Ethanol costs and energy balances.

Feedstock	Cost (US$ per gallon)	Energy balance (renewable output to fossil input)
Sugarcane, Brazil		10.2 (*18*)
2006, without import tax	0.81 (*17*)	
2006, with U.S. import tax	1.35 (*9*, *17*)	
Sugar beet, Europe, 2003	2.89 (*17*)	2.1 (*19*)
Corn, U.S., 2006	1.03 (*17*)	1.4 (*9*, *11*)
Cellulose ethanol, U.S.		10.0 (*11*)
Achieved in 2006	2.25 (*11*)	
Target for 2012	1.07 (*11*)	

nes (30% or less with modern flexible fuel vehicle technologies) (9). The summer wholesale price of gasoline in the United States is about $1.9 per gallon; the corn ethanol wholesale price is around US$2.5 per gallon (*10*). Cellulose ethanol is a promising option in the long term, but is not being produced on a commercial scale. The longer-term target is as low as 60 cents per gallon, but this will require major advances in producing, collecting, and converting biomass. A more realistic research target is to reduce the cost of production to US$1.07 per gallon until 2012 (*11*).

The development of other biomass-derived fuels in Brazil or elsewhere could benefit from such insights. Promising candidates along those lines are the following.

1) The production of ethanol from cellulosic materials, which still requires considerable R&D effort before reaching the production stage. If the technology for such conversion is firmly established, it would open enormous opportunities for the use of all kinds of wood and other biomass feedstocks for ethanol production.

2) The enhanced use of biogas produced from microbial conversion in landfills of municipal solid wastes, wastewater, industrial effluents, and manure wastes will abate a considerable share of greenhouse gases that would be released to the atmosphere, replacing also fossil fuels for heat and electricity production.

3) The use of planted forests for the production of electricity either by direct combustion or by gasification and use of highly efficient gas turbines will also replace efficiently coal, natural gas, oil, and even nuclear sources. Reforested wood can also reduce the need for deforested fuel wood, controlling efficiently releases of greenhouse gases through market-friendly initiatives.

The ethanol program in Brazil was based on indigenous technology (both in the industrial and agricultural areas) and, in contrast to wind and solar photovoltaics, does not depend on imports, and the technology can be transferred to other developing countries.

Until breakthrough technologies become commercially viable, an alternative already exists. Many developing countries have suitable conditions to expand and replicate the Brazilian sugarcane program, supplying the world's gasoline motor vehicles with a renewable, efficient fuel.

References and Notes

1. United Nations, Report of the World Commission on Environment and Development, United Nations General Assembly, 96th plenary meeting, 11 December 1987, Document A/RES/42/187, available at www.un.org/documents/ga/res/42/ares42-187.htm
2. British Petroleum, BP Statistical Review of World Energy, available at www.bp.com/liveassets/bp_internet/globalbp/globalbp_uk_english/publications/energy_reviews_2006/STAGING/local_assets/downloads/spreadsheets/statistical_review_full_report_workbook_2006.xls.
3. J. Goldemberg, "The case for renewable energies" (background paper for the International Conference for Renewable Energies, Bonn 2004), available at www.renewables2004.de/pdf/tbp/TBP01-rationale.pdf. On jobs see also (*12*).
4. REN21, Global Status Report 2006 Update (Renewable Energy Policy Network for the 21st Century, 2006); available at www.ren21.net/pdf/RE_GSR_2006_Update.pdf.
5. FAO, FAOSTAT (United Nations Food and Agriculture Organization, 2006), available at http://faostat.fao.org/default.aspx.
6. J. Goldemberg, S. T. Coelho, O. Lucon, P. M. Nastari, *Biomass Bioenergy* **26**, 301 (2004).
7. J. Goldemberg, S. T. Coelho, O. Lucon, *Energy Policy* **32**, 1141 (2004).
8. J. G. Da Silva, G. E. Serra, J. R. Moreira, J. C. Gonçalves, J. Goldemberg, *Science* **201**, 903 (1978).
9. S. T. Coelho, J. Goldemberg, O. Lucon, P. Guardabassi, *Development* **10**, 26 (2006). On ethanol duties see also (*13*).
10. J. R. Healey, "Ethanol shortage could up gas prices," *USA Today*, 30 March 2006, available at www.usatoday.com/money/industries/energy/2006-03-30-ethanol-gas-prices_x.htm.
11. M. Pacheco, U.S. Senate Full Committee Hearing: Renewable Fuel Standards (National Renewable Energy Laboratory, National Bioenergy Center, 19 June 2006); available at http://energy.senate.gov/public/index.cfm?isPrint=true&FuseAction=Hearings.Testimony&Hearing_ID=1565&Witness_ID=4427. On corn ethanol, see also (*14*).
12. World Bank, "How the World Bank's energy framework sells the climate and poor people short" (World Bank, September 2006); available at www.nirs.org/climate/background/energyreportfinal91806.pdf.
13. A. Elobeid, S. Tokgoz, Removal of U.S. Ethanol Domestic and Trade Distortions: Impact on U.S. and Brazilian Ethanol Markets (Working Paper 06-WP 427, October 2006); available at www.card.iastate.edu/publications/DBS/PDFFiles/06wp427.pdf.
14. U.S. Department of Agriculture, Estimating the Net Energy Balance of Corn Ethanol, M. Shapouri, J. A. Duffield, M. S. Graboski (U.S. Department of Agriculture, Economic Research Service, Office of Energy, Agricultural Economic Report No. 721); available at www.ers.usda.gov/publications/aer721/AER721.PDF.
15. UNDP, UNDESA, WEC, World Energy Assessment Overview 2004 Update (United Nations Development Program, United Nations Department of Economic and Social Affairs, World Energy Council, 2005), available at www.undp.org/energy/weaover2004.htm.
16. IEA, Key World Energy Statistics (International Energy Agency, 2006); available at www.iea.org/w/bookshop/add.aspx?id=144.
17. USDA, *The Economic Feasibility of Ethanol Production from Sugar in the United States* (United States Department of Agriculture, 2006).
18. I. C. Macedo, "Greenhouse gas emissions and energy balances in bio-ethanol production and use in Brazil" available at www.unica.com.br/i_pages/files/gee3.pdf.
19. J. Woods, A. Bauen, "Technology status review and carbon abatement potential of renewable transport fuels in the UK" (United Kingdom Department of Transport and Industry Report BAU/00785/REP URN 03/987); available at www.dti.gov.uk/files/file15003.pdf.

20. I thank O. Lucon and J. R. Moreira for useful discussions and contributions.

10.1126/science.1137013

PERSPECTIVE

Renewable Energy Sources and the Realities of Setting an Energy Agenda

Janez Potočnik

The European Commission has been devoting considerable attention to energy issues. This Perspective describes recent progress in Europe toward achieving goals for renewable energy use, and the role that technology can play, as well as the new Strategic Energy Package.

Energy is undoubtedly moving up the political agenda as old issues come to be addressed urgently. If last year's threats to European gas supplies during the dispute between Russia and Ukraine did not show the immediacy of the challenges such as energy supply, then the report toward the end of last year by Sir Nicholas Stern (*1*) on the economics of climate change must surely have rung a warning bell.

The European Commission has been devoting considerable attention to energy issues for some time now. We were leaders in the process that brought about the Kyoto Protocol and have developed the first large-scale emissions trading scheme in the world. In March 2006, we published a Green Paper on energy (*2*), which we have now, at the beginning of 2007, followed up with a strategic energy package (*3*) addressing energy policy in general and also outlining future European policy on various specific elements.

One of these specific elements will be the elaboration at the European level of a Strategic Energy Technology Plan (*4*). Research and technology will undoubtedly be crucial to cracking the energy and climate change nut. A recent study published by the European Commission (Fig. 1) (*5*) shows that, if existing trends

European Commission, 1049 Bruxelles, Belgium.

continue, by 2050 CO_2 emissions will be unsustainably high: 900 to 1000 parts per million by volume, that is, well above what is considered an acceptable range for stabilization. Without determined action, energy demand will double and electricity demand will quadruple, resulting in an 80% increase in CO_2 emissions. However technological development coupled with strong carbon constraint policies can limit this impact, with world emissions stable between 2015 and 2030 and decreasing thereafter. In this "carbon constraint" case, half of the total building stock would be made of low-energy buildings, and more than half of the vehicles would have low or very low emissions, a clear example of how technological development will contribute to our energy and environmental policy objectives.

The strategic energy package sets a target of 20% of Europe's energy coming from renewable sources by 2020. If successful, this would mean that by 2020 the European Union (EU) would use about 13% less energy than today, saving €100 billion and around 780 metric tons of CO_2 each year. For this to be realistic, significant strides need to be made technologically speaking. Today renewable energy is on the whole costly and intermittent. Even if we are looking to maintain a mix of sources of energy, a cloudy windless day rules out generation from solar and wind power. And yet, on a bright windy day, energy may go unused, because it cannot be stored easily. Reliability and continuity are basic requirements if renewable sources of energy are to be seen as viable alternatives to oil, gas, and coal. Research and technological development are already bringing us closer to solutions in this field, through improving fuel cells or redesigning electricity grids to deal with more decentralized power generation.

Fig. 1.

We believe that renewables have the potential to provide around a third of EU electricity by 2020 (3). Current statistics indicate that this is not an unreasonable goal. Wind power currently provides roughly 20% of electricity needs in Denmark, as well as 8% in Spain and 6% in Germany. If other Member States matched the levels that Sweden, Germany, and Austria have attained in geothermal heat pumps and solar heating, the share of renewable energy in heating and cooling would jump by 50%. As for biofuels, Sweden has already achieved a market share of 4% of the petrol market for bioethanol, and Germany is the world leader for biodiesel, with 6% of the diesel market. Biofuels could account for as much as 14% of transport fuels by 2020 (3). The European public is also clearly in favor of advancing renewable sources of energy, with a recent opinion poll (6) showing approval ratings for such energy ranging between 55% and 80%.

The European Commission has certainly taken this on board in its new research funding program, the Seventh Framework Programme (7). Within the energy theme of the cooperation program, which will focus on nonearbon or reduced-carbon sources of energy, emphasis will be given to renewable electricity generation, renewable fuel production, hydrogen and fuel cells, CO_2 capture and storage technologies, smart energy networks, energy efficiency and savings, nuclear fission safety and waste management, the development of fusion energy, and knowledge for energy policy-making. The Seventh Framework Programme increases the annual funding available to energy research at the European level to €886 million a year, compared to €574 million a year in the previous program. But this is not enough: more combined effort is needed. In some areas, we have moved toward common research agendas at the European level through the creation of European technology platforms (8). Several exist in the energy field, including for hydrogen and fuel cells, photovoltaics, zero-emission fossil-fuel power plants, and smart grids. Nonetheless, we have seen investment in energy research being reduced in national budgets over the past 20 years or so. And the research that is carried out is more often than not done in a fragmented, uncoordinated way, leading to duplication in some areas and other important aspects being underfunded or ignored. This is the raison d'etre of the Strategic Energy Technology Plan, which will, once agreed on, provide a basis for all energy technology efforts in Europe, overcoming the lack of coherence that has unfortunately been present to a greater or lesser extent in the research programs at the national and European levels up to now.

During the first half of 2007, the Commission will consult intensively with all those that have a role to play in such a strategic plan. On the basis of these consultations, a text will be drawn up toward the middle of the year, upon which the research community, among others, will be invited to give its comments. It is important that the creation of the Strategic Energy Technology Plan is a collaborative bottom-up process if it is to have any chance of achieving its stated objective of being a reference point for future European Union activities in this area.

Since my appointment as European Science and Research Commissioner in November 2004, I have insisted on the importance of science and research as the key to solving many of the challenges that we face. I can think of no better illustration of this approach than the issue of energy. Here, we have various requirements in front of us: finding secure and sustainable sources of energy that support our economic growth and competitiveness without damaging our environment. The answer to reconciling these requirements lies in knowing more and being better. We have a chance to work together to develop solutions to the problems of climate change and energy supply that not only ensure our future economic development, but give European scientists and companies the opportunity to be (or remain) at the cutting edge of technological development. It is crucially important that we take this opportunity and make it work.

References and Notes

1. Stern Review on the Economics of Climate Change, 2006. www.hm-treasury.gov.uk/independent_reviews/stern_review_economics_climate_change/stern_review_index.cfm
2. Commission of the European Communities, *Green Paper A European Strategy for Sustainable, Competitive and Secure Energy*, SEC(2006) 317, http://ec.europa.eu/energy/green-paper-energy/doc/2006_03_08_gp_document_en.pdf
3. Commission of the European Communities, *Communication from the Commission to the European Council and the European Parliament: An Energy Policy for Europe*, COM(2007) final, http://ec.europa.eu/energy/energy_policy/doc/01_energy_policy_for_europe_en.pdf
4. European Commission, *Toward A Strategic Energy Technology Plan*, COM(2007) 847, ec.europa.eu/energy/energy_policy/doc/19_strategic_energy_technology_plan_en.pdf.
5. European Commission, *World Energy Technology Outlook—2050*, 2006, http://ec.europa.eu/research/energy/pdf/weto-h2_en.pdf
6. European Commission, *Energy Technologies: Knowledge, Perception, Measures*, 2006, http://ec.europa.eu/research/energy/pdf/energy_tech_eurobarometer_en.pdf
7. European Commission, *Official Journal*, document L412 (2006), http://eur-lex.europa.eu/LexUriServ.do?uri=OJ:L:2006:412:0001:EN:HTML
8. For further information, see http://cordis.europa.eu/technology-platforms/individual_en.html.
9. The author is the European Commissioner for Science and Research.

10.1126/science.1139086

PERSPECTIVE

Preparing to Capture Carbon

Daniel P. Schrag

Carbon sequestration from large sources of fossil fuel combustion, particularly coal, is an essential component of any serious plan to avoid catastrophic impacts of human-induced climate change. Scientific and economic challenges still exist, but none are serious enough to suggest that carbon capture and storage will not work at the scale required to offset trillions of tons of carbon dioxide emissions over the next century. The challenge is whether the technology will be ready when society decides that it is time to get going.

Strategies to lower carbon dioxide (CO_2) emissions to mitigate climate change come in three flavors: reducing the amount of energy the world uses, either through more efficient technology or through changes in lifestyles and behaviors, expanding the use of energy sources that do not add CO_2 to the atmosphere, and capturing the CO_2 from places where we do use fossil fuels and then storing it in geologic repositories, a process known as carbon sequestration. A survey of energy options makes clear that none of these is a silver bullet. The world's energy system is too immense, the thirst for more and more energy around the world too deep, and our dependence on fossil fuels too strong. All three strategies are essential, but the one we are furthest from realizing is carbon sequestration.

The crucial need for carbon sequestration can be explained with one word: coal. Coal produces the most CO_2 per unit energy of all fossil fuels, nearly twice as much as natural gas. And unlike petroleum and natural gas, which are predicted to decline in total production well before the middle of the century, there is enough coal to last for centuries, at least at current rates of use, and that makes it cheap relative to almost every other source of energy (Table 1). Today, coal and petroleum each account for roughly 40% of global CO_2 emissions. But by the end of the century coal could account for more than 80%. Even with huge improvements in efficiency and phenomenal rates of growth in nuclear, solar, wind, and biomass energy sources, the world will still rely heavily on coal, especially the five countries that hold 75% of world reserves: the United States, Russia, China, India, and Australia (1).

As a technological strategy, carbon sequestration need not apply only to coal plants; indeed, any point source of CO_2 can be sequestered, including biomass combustion, which would result in negative emissions. Carbon sequestration also refers to enhanced biological uptake through reforestation or fertilization of marine phytoplankton. But the potential to enhance biological uptake of carbon pales in comparison to coal emissions, ever more so as India, China, and the United States expand their stock of coal-fired power plants. So developing and deploying the technologies to use coal without releasing CO_2 to the atmosphere may well be the most critical challenge we face, at least for the next 100 years, until the possibility of an affordable and completely nonfossil energy system can be realized.

If carbon sequestration from coal combustion is essential to mitigate the worst impacts of global warming, what stands in the way of its broad implementation, both in the United States and around the world? With limited coal reserves, countries in the European Union have chosen to emphasize climate mitigation strategies that focus on energy efficiency, renewable sources, and nuclear power. Of the major coal producers, Russia, China, and India have been unwilling to sacrifice short-term economic growth, although Chinese coal gasification efforts, which many see as a step toward sequestration capacity, are more advanced than current U.S. policies. In the United States, there are scientific and economic questions that must be answered before large-scale deployment can be achieved. But none of these is critical enough to suggest that carbon sequestration cannot be done. The real obstacle is political will, which may require more dramatic public reaction to climate change impacts before carbon sequestration becomes a requirement for burning coal. In the meantime, there are critical steps that can be taken that will prepare us for the moment when that political will finally arrives.

The scientific questions about carbon sequestration are primarily associated with concerns about the reliability of storage of vast quantities of CO_2 in underground repositories. Will the CO_2 escape? The good news is that the reservoirs do not have to store CO_2 forever, just long enough to allow the natural carbon cycle to reduce the atmospheric CO_2 to near pre-industrial levels. The ocean contains 50 times as much carbon as the atmosphere, mostly in the deep ocean, which has yet to equilibrate with the CO_2 from fossil fuel combustion. Over the time scale of mixing of the deep ocean, roughly 1000 to 2000 years, natural uptake of CO_2 by the ocean, combined with dissolution of marine carbonate, will absorb 90% of the carbon released by human activities. As long as the geologic storage of CO_2 can prevent substantial leakage over the next few millennia, the carbon cycle can handle it.

Our current understanding of CO_2 injection in sedimentary reservoirs on land suggests that leakage rates are likely to be very low (2). Despite many years of experience with injection of CO_2 for enhanced oil recovery, few studies have accurately measured the leakage rates over time intervals long enough to be certain that the CO_2 will stay put even for the next few centuries. In most of the geological settings under consideration, such as deep saline aquifers or old oil and gas fields, CO_2 exists as a supercritical fluid with roughly half the density of water. CO_2 is trapped by low-permeability cap rocks and by capillary forces, but can escape if sedimentary formations are compromised by fractures, faults, or old drill holes. The handful of test sites around the world each inject roughly 1 million tons of CO_2 per year, a tiny amount compared to the need for as much as 10 billion tons per year by the middle of the century. An important question is whether leakage rates will rise as more and more CO_2 is injected and the reservoirs fill. It seems likely that many geological settings will provide adequate storage, but the data to demonstrate this do not yet exist. A more expansive program aimed at monitoring underground CO_2 injections in a wide variety of geological settings is essential.

A recent proposal identified a leak-proof approach to storage by injecting CO_2 in sediment below the sea floor (3), which avoids the hazards of

Table 1. Carbon content in gigatons (Gt) of fossil fuel, proven reserves and annual production (2005) (6).

Country/region	Coal		Petroleum		Natural gas	
	Reserves	Production	Reserves	Production	Reserves	Production
United States	184.0	0.64	3.6	0.30	3.0	0.29
Russia	117.1	0.15	9.0	0.42	26.2	0.33
China	85.4	1.24	1.9	0.16	1.3	0.03
India	69.0	0.22	0.7	0.03	0.6	0.02
Australia	58.6	0.23	0.5	0.02	1.4	0.02
Middle East	0.3	0.00	90.2	1.11	39.4	0.16
Total world	678.2	3.23	145.8	3.59	98.4	1.51

Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA. E-mail: schrag@eps.harvard.edu

direct ocean injection, including impacts on ocean ecology. In this case, CO_2 would stay separate from the ocean, because it exists in the sediment at high pressure and low temperature as a dense liquid or combined with pore fluid as solid hydrate. Despite higher possible costs, this approach may be important for coastal locations, which are far from appropriate sedimentary basins, and may also avoid expensive monitoring efforts if leakage from terrestrial settings is found to be a major problem.

In terms of capacity, the requirements are indeed vast. Conservative estimates of reservoir needs over the century are more than 1 trillion tons of CO_2, and might exceed twice that much. This far exceeds the capacity of oil and gas fields, which will be among the first targets for sequestration projects because of additional revenues from enhanced oil recovery. Fortunately, the capacity of deep saline aquifers and deep-sea sediments is more than enough to handle centuries of world coal emissions (*3*, *4*). This means that the locations first used to store CO_2 underground may not be the ones used by the middle of the century as sequestration efforts expand. It suggests that a broad research program must be encouraged that focuses not just on what will be done in the next few decades, but also on approaches that will be needed at the scale when all coal emissions will be captured.

Additional questions surround the more expensive part of carbon sequestration, the capture of CO_2 from a coal-fired power plant. Conventional pulverized coal plants burn coal in air, producing a low-pressure effluent composed of 80% nitrogen, 12% CO_2, and 8% water. CO_2 can be scrubbed from the nitrogen using amine solids or other chemicals, and then extracted and compressed for injection into storage locations. This uses energy, roughly 30% of the energy from the coal combustion in the first place (*4*), and may raise the generating cost of electricity from coal by 50% (*5*), although these estimates are uncertain given that there is not yet a coal plant that practices carbon sequestration. Pulverized coal plants can also be retrofit to allow for combustion of coal in pure oxygen, although the separation of oxygen from air is similarly energy intensive, and the modifications to the plant would be substantial and likely just as costly (*4*).

Gasification of coal, which involves heating and adding pure oxygen to make a mixture of carbon monoxide and hydrogen, can be used either for synthesis of liquid fuels or for electricity. These plants can be designed to produce concentrated streams of pressurized CO_2, often referred to as "capture-ready," although this also comes at a high cost. Much attention has been given to coal gasification as a means for promoting carbon sequestration because studies suggest that the costs are lower than retrofitting an existing pulverized coal plant (*4*). However, experience with gasification plants is limited; there are only two such plants in the United States, and neither is capture ready. More encouragement of coal gasification technology is important to discover whether the promises of lower sequestration costs can be realized. But regardless of the emphasis on such advanced coal plants, the world's existing arsenal of pulverized coal plants (excluding the 150 new pulverized coal plants that are currently in the permitting process in the United States) produce roughly 8 billion tons of CO_2 per year, more than any responsible climate change policy can accommodate. Thus, the investment in advanced coal gasification plants must be matched by an effort to optimize our ability to capture the CO_2 from existing pulverized coal plants.

Compared with the cost of most renewable energy sources, increasing the cost of electricity from coal by 50% to add sequestration seems like a bargain. When one includes the distribution and delivery charges, electric bills of most consumers would rise only 20% or so. So why is this not a higher priority in climate change legislation? Most legal approaches to climate change mitigation have focused on market mechanisms, primarily cap and trade programs. A problem is that the cap in Europe and any of the caps under discussion in the U.S. Congress yield a price on carbon that is well below the cost of capture and storage. Even if the caps were lowered, power companies might hesitate to invest in the infrastructure required for sequestration because of volatility in the price of carbon. Thus, it seems that another mechanism is required, at least to get carbon sequestration projects started.

And there are many other questions. Who will certify a storage site as appropriate? How will the capacity be determined? Who will be responsible if CO_2 leaks? How will we safeguard against cheating? It is clear that governments need to play some role in CO_2 storage, just as they do in other forms of waste disposal, but the exact details of a policy are unlikely to be decided in the near future, long before carbon sequestration becomes normal practice. But the uncertainty about these and other issues contributes to a general cloudiness that discourages industry from making investments toward sequestration efforts.

Despite these obstacles, a variety of carbon sequestration activities are proceeding. Regional partnerships have been established in the United States, supported by the U.S. Department of Energy (DOE), to study the possibilities for sequestration around the country. In 2003, President Bush announced a commitment to FutureGen, a DOE project to build a zero-emission coal gasification plant that would capture and store all the CO_2 it produced. FutureGen is an exciting step forward, but a single coal gasification plant that demonstrates carbon sequestration is unlikely to convince the world that carbon sequestration is the right strategy to reduce CO_2 emissions. Moreover, a power plant operated by the government may fail to convince power companies that the costs of sequestration are well determined.

Luckily, FutureGen has competitors. British Petroleum (BP), in cooperation with General Electric, plans to build two electricity-generating plants, one in Scotland and one in California, that would sequester CO_2 with enhanced oil recovery. Xcel Energy has also made a commitment to build a coal gasification plant with sequestration. And more projects may soon be announced as companies begin to view legislation controlling CO_2 emissions as a political inevitability.

Given the current questions about sequestration technology, the current economic realities that make it unlikely that many companies will invest in sequestration over a sustained period, and the political realities that make it unlikely we will see in the next few years a price on carbon high enough to force sequestration from coal, what can government do to make sure that carbon sequestration is ready when we need it? Whatever the path, it is time to get going, not just with small test projects but with full-scale industrial experiments. The announcements by BP and Xcel Energy are encouraging because the world needs many such sequestration projects operating at different locations, with a handful of capture strategies and a wider variety of geological settings for storage. The U.S. government can encourage these efforts, and sponsor additional ones, making sure that there are 10 to 20 large sequestration projects operating for the next decade so that any problems that do arise with capture or storage can be identified. By creating a competitive bidding process for long-term sequestration contracts, the United States can ensure that the most cost-efficient strategies will be used while testing a variety of capture and storage options including retrofitting older pulverized coal plants. The United States and the world need carbon sequestration—not right now, but soon and at an enormous scale. Our challenge today is to ensure that the technology is ready when serious political action on climate change is finally taken.

References and Notes

1. British Petroleum, BP Statistical Review of World Energy (2006); available at www.bp.com/productlanding.do?categoryId=6842&contentId=7021190.
2. J. Bradshaw, C. Boreham, F. La Pedalina, Storage retention time of CO_2 in sedimentary basins: Examples from petroleum systems, Report of the Greenhouse Gas Technologies Cooperative Research Centre, Canberra (2004), available at www.co2crc.com.au/PUBFILES/STOR040505GHGT7_Bradshaw_Boreham_LaPedalina.pdf.
3. K. Z. House, D. P. Schrag, C. F. Harvey, K. S. Lackner, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12291 (2006).
4. IPCC Special Report on Carbon Dioxide Capture and Storage (Intergovernmental Panel on Climate Change, 2005); available at www.ipcc.ch/activity/srccs/index.htm.
5. S. Anderson, R. Newell, *Annu. Rev. Environ. Resour.* **29**, 109 (2004).
6. Data on proven reserves and production are from (*1*). Carbon content was calculated assuming 25.4 tons (1 ton = 10^3 kg) of carbon per terajoule (TJ) of coal; 19.9 tons of carbon per TJ of petroleum; and 14.4 tons of carbon per TJ of natural gas. Differences between anthracite, bituminous, lignite, and sub-bituminous coal were not included.

7. The author benefited from discussions with K. House and I. Holdren.

10.1126/science.1137632

Enceladus: Cosmic Graffiti Artist Caught in the Act

Anne Verbiscer,[1]* Richard French,[2] Mark Showalter,[3] Paul Helfenstein[4]

Saturn's geologically active moon Enceladus is energized by the gravitational giant who produces Mount Etna's volcanic fires. The moon's south polar plume eruptions give rise to the vast, tenuous E ring that enshrouds at least 11 satellites. Despite its small size, Enceladus is even more deserving of its gigantic namesake. Here, we report Hubble Space Telescope (HST) observations of Saturn's satellites at true opposition that show how material originating from Enceladus alters the appearance of its E-ring neighbors.

On the night of 13 January 2005, a rare perfect alignment of the Sun, Earth, and Saturn enabled the measurement of the albedos of the saturnian satellites at the smallest solar phase (Earth-Sun-Saturn) angle ($\alpha \geq 0.01°$) attainable at Saturn's heliocentric distance. Previously, their geometric albedos were poorly determined and could be estimated only by extrapolating photometric models to zero phase. Geometric albedo, p, is the ratio of a satellite's reflectance at $\alpha = 0°$ to that of a perfectly diffusing disk viewed at the same position and apparent size. All albedos measured at true opposition are considerably higher than previous estimates (table S1). For example, p is 1.23 for Tethys versus 0.80 previously, and p is 1.00 for Dione versus 0.55 previously (2). With the exception of Janus (p = 0.71) and Epimetheus (p = 0.73), which orbit interior to the E ring, the albedos of all satellites measured here approximate or exceed unity (3). A strong correlation exists (Fig. 1) between p and the pole-on radial reflectance profile of the E ring as a function of orbital radius. Calypso and Helene, Lagrangian satellites of Tethys and Dione, respectively, are visible within HST's field of view. Both also have $p > 1$; however, these values are uncertain because of low signal-to-noise ratios and large corrections for charge transfer efficiency. In addition, Helene's shape is not well represented by a simple triaxial ellipsoid.

The extraordinarily high albedo of Enceladus is reasonable for a geologically active body undergoing continuous resurfacing. The fact that $p \geq 1$ for all embedded satellites is, however, striking, because the other bodies have ancient, inactive surfaces. Most icy moons in the outer solar system have p values between 0.2 and 0.4 (2); exceptions include satellites with relatively young surfaces such as Jupiter's Europa [p = 0.9 (4)] and Neptune's Triton [p = 0.8 (5)], which also has active geysers.

We propose that interactions with E-ring particles, ultimately from Enceladus' plumes, produce the high albedos on Tethys, Mimas, Dione, and Rhea. After ejection from Enceladus' south pole, nongravitational forces excite the eccentricities of E-ring grains (6). These particles collide with the embedded satellite surfaces at high relative velocities (6). Most of the ejecta resulting from this sandblasting falls back onto the satellite surfaces, coating them with clean, icy microstructures suited to enhancing the reflectance at opposition (7).

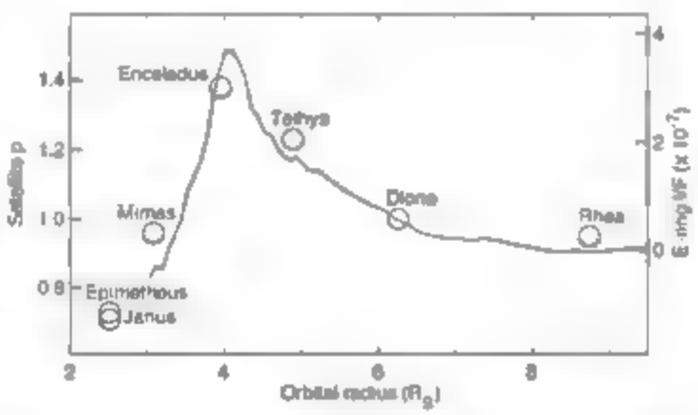


Fig. 1. The mean visual geometric albedo, p (left vertical axis), of the embedded satellites Mimas (0.96), Enceladus (1.38), Tethys (1.23), Dione (1.00), and Rhea (0.95) versus radial distance from Saturn (R_S) mimics the pole-on reflectance (I/F) profile (solid line and right vertical axis) of the E ring. HST's Wide Field and Planetary Camera 2 (WFPC2) observed both the 1995 ring plane crossings (*10*) and the satellites at true opposition with the same filter (F555W).

Although a more complete understanding of the regolith physical properties is ultimately expected from the analyses of full solar phase curves (reflectance versus α for $0° < \alpha < 180°$), a preliminary analysis of the phase curves near opposition ($\alpha < 6.4°$) using the Hapke (*8*) model suggests that differences in the opposition effect (OE) among the embedded satellites may indeed be related to the degree to which they encounter E-ring grains. OE, the dramatic, nonlinear increase in reflectance as $\alpha \to 0°$, is characterized by its amplitude and angular width, both of which depend on the microphysical structure and density of regolith grains (*9*). The OE amplitudes and widths of the embedded satellites' phase curves correlate with position in the E-ring (fig. S1). Three groupings of satellites emerge with similar OE amplitudes and widths: Tethys and Enceladus, Mimas, Dione and Rhea, and Janus and Epimetheus. In models that successfully reproduce the E-ring brightness profile, Tethys and Enceladus experience substantially higher impact rates from E-ring particles than do Mimas, Dione, and Rhea (*9*). E-ring particles do not impact Janus and Epimetheus.

At least four additional small satellites orbit within the E ring, two between Mimas and Enceladus and two more Lagrangian companions of Tethys and Dione. Although measurements must await the acquisition of new Cassini images, we predict that [illegible] of these moons as well. Inevitably, material from Enceladus impacts all satellites orbiting within the E ring, enhancing their albedos at the hands of a diminutive graffiti artist.

References and Notes

1. C. C. Porco *et al.*, *Science* **311**, 1393 (2006).
2. J. A. Burns, in *Satellites*, J. A. Burns, M. S. Matthews, Eds. (Univ. of Arizona Press, Tucson, AZ, 1986), pp. 1–38.
3. Conservation of energy allows $p > 1$ if a satellite is strongly backscattering or has a strong OE.
4. D. L. Domingue *et al.*, *Icarus* **90**, 30 (1991).
5. M. Hicks, B. Buratti, *Icarus* **171**, 210 (2004).
6. D. P. Hamilton, J. A. Burns, *Science* **264**, 550 (1994).
7. Y. Shkuratov, P. Helfenstein, *Icarus* **152**, 96 (2001).
8. B. Hapke, *Icarus* **157**, 523 (2002).
9. A. Juhász, M. Horányi, *J. Geophys. Res.* **107**, 1066 (2002).
10. P. D. Nicholson *et al.*, *Science* **272**, 509 (1996).
11. Based on observations with NASA/European Space Agency's HST, obtained at the Space Telescope Science Institute. P.H. acknowledges support from NASA's Planetary Geology and Geophysics program.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5813/815/DC1
Fig. S1
Tables S1 and S2
References

5 September 2006; accepted 7 November 2006
10.1126/science.1134681

[1]Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA. [2]Department of Astronomy, Wellesley College, Wellesley, MA 02481, USA. [3]Search for Extraterrestrial Intelligence Institute, Mountain View, CA 94043, USA. [4]Department of Astronomy, Cornell University, Ithaca, NY 14853, USA.

*To whom correspondence should be addressed. E-mail: verbiscer@virginia.edu

Direct Measurements of the Convective Recycling of the Upper Troposphere

Timothy H. Bertram,[1] Anne E. Perring,[2] Paul J. Wooldridge,[1] John D. Crounse,[2] Alan J. Kwan,[3] Paul O. Wennberg,[3,4] Eric Scheuer,[5] Jack Dibb,[5] Melody Avery,[6] Glen Sachse,[6] Stephanie A. Vay,[6] James H. Crawford,[6] Cameron S. McNaughton,[7] Antony Clarke,[7] Kenneth E. Pickering,[8,9] Henry Fuelberg,[10] Greg Huey,[11] Donald R. Blake,[12] Hanwant B. Singh,[13] Samuel R. Hall,[14] Richard E. Shetter,[14] Alan Fried,[14] Brian G. Heikes,[15] Ronald C. Cohen[1,16]*

We present a statistical representation of the aggregate effects of deep convection on the chemistry and dynamics of the upper troposphere (UT) based on direct aircraft observations of the chemical composition of the UT over the eastern United States and Canada during summer. These measurements provide unique observational constraints on the chemistry occurring downwind of convection and the rate at which air in the UT is recycled. These results provide quantitative measures that can be used to evaluate global climate and chemistry models.

Deep convection is a highly efficient mechanism for the vertical transport of air from near Earth's surface (0 to 2 km) to the UT (6 to 12 km) (*1*–*5*). Typical convective storms have spatial scales of tens of kilometers and vertical velocities as large as 15 m s^{-1} (*6*), making their local influence in the UT extremely strong. The rapid upward flow is balanced by downdrafts within the convective storms and much slower descending flow that occurs over a larger spatial scale (*7*). Convection is also associated with lightning, an important source of NO_x ($NO_x = NO + NO_2$) in the UT (*8*, *9*). The source strength and spatial distribution of lightning NO_x emissions are not well known, with estimates ranging from 2 to 20 Tg N year^{-1} for the global average (*10*), compared to 25 Tg N year^{-1} from fossil fuel combustion (*11*). Although there have been a number of case studies of the chemical effects of individual storms (*12*), studies of the aggregate effects of convection on the chemical composition and radiative forcing of the UT have been largely the province of modeling and theory (*13*, *14*). Here, we describe measurements that provide a direct link between an observable property and the ensemble of convective events.

The chemical and radiative consequences of convection and lightning are known to be large (*2*, *15*, *16*). Upper tropospheric O_3, either transported directly from the boundary layer via convection or formed in situ after detrainment of convectively lofted O_3 precursors (NO_x, odd hydrogen radicals (HO_x), and hydrocarbons) in the outflow region, directly affects climate through a positive radiative forcing (*15*). Additionally, deep convection accounts for a substantial fraction of the net flux of moisture from near Earth's surface to the UT (*17*) (Fig. 1). Thus, the rate at which the UT is turned over by convection has important implications for the hydrological cycle and the magnitude of the water vapor feedback on global temperature (*18*).

We describe a method for calculating, from in situ measurements of the chemical composition of the UT, the length of time that an air mass spends in the UT after convection, and we discuss the chemistry occurring in the outflow region as a function of time since convection. We use measurements of NO_2 (*19*, *20*) (NO_x is calculated from NO_2, O_3, HO_2, and photolysis rates), HNO_3 (*21*, *22*), OH and HO_2 (*23*), O_3, aerosol number density (*24*), SO_2, net flux [from which photolysis rate coefficients (J_x) for NO_2 (J_{NO_2}), HNO_3 (J_{HNO_3}), and many other species are calculated] (*25*), CO (*26*), and CO_2 (*27*) obtained during the Intercontinental Chemical Transport Experiment North America (INTEX-NA) aboard the NASA DC-8 (*28*). Measurements were made at altitudes between the surface and 12.5 km over a wide area of the United States and Canada, west of 40°W and between 30° and 50°N. There were a large number of vertical profiles, allowing a reasonably unbiased statistical sampling of air over this region during July and August 2004.

We use the deviation of the observed NO_y:HNO_3 ratio from steady state as an indicator of convective influence. The NO_y:HNO_3 ratio is reset to near unity in most convection as a result of preferential wet scavenging of HNO_3 relative to NO_x (i.e., the solubility of HNO_3 is 10^5 times that of NO_x) (*29*). Further, lightning-produced NO_x, often coincident with convection, markedly enhances NO_x in the outflow region. The coupling of these processes makes the NO_y:HNO_3 ratio in the UT an effective indicator of convective influence, where NO_y:$HNO_3 >> 1$ is indicative of recent cloud outflow (*30*, *31*). In the days after convection, the ratio decays toward steady state, providing a chemical clock that marks the length of time that an air mass has spent in the UT after convection (*32*). Previous studies have used species that have no UT source (e.g., CH_3I) (*33*) or alternative chemical ratios to provide estimates of age of air in the UT (*34*, *35*). Our study is unique because of the availability of

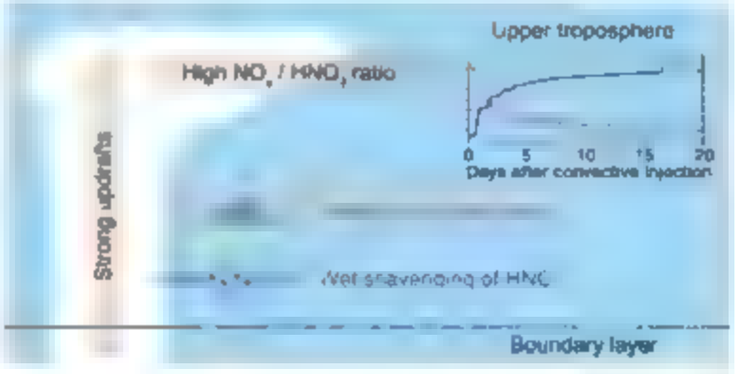


Fig. 1. In most convection, air from near Earth's surface is rapidly transported upward and detrained into the UT. In this process, nitric acid (highly soluble) is efficiently scavenged while NO_y (insoluble) remains. NO_y is elevated by concurrent lightning NO production, resulting in high NO_y:HNO_3 ratios in the convective outflow region. After detrainment into the UT, NO_y is converted to HNO_3 by OH during the day and through reaction with NO_3, followed by hydrolysis of the N_2O_5 product, at night. The chemical evolution of the NO_y:HNO_3 ratio provides a unique indicator of the length of time that a sampled air mass has been in the UT after convection.

[1]Department of Chemistry, University of California, Berkeley, CA 94720, USA. [2]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA. [3]Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA. [4]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA. [5]Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH 03824, USA. [6]NASA Langley Research Center, Hampton, VA 23681, USA. [7]School of Ocean and Earth Science Technology, University of Hawaii at Manoa, Honolulu, HI 96822, USA. [8]Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD 20742, USA. [9]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. [10]Department of Meteorology, Florida State University, Tallahassee, FL 32306, USA. [11]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA. [12]Department of Chemistry, University of California, Irvine, CA 92697, USA. [13]NASA Ames Research Center, Moffett Field, CA 94035, USA. [14]National Center for Atmospheric Research, Boulder, CO 80305, USA. [15]Graduate School of Oceanography, University of Rhode Island, Narragansett, RI 02882, USA. [16]Department of Earth and Planetary Science, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: cohen@rchem.berkeley.edu

NO_2, OH, and HNO_3 measurements with high time resolution, which allows us to build a much more extensive data set than in previous studies. After the initial turbulent mixing in the near field of the convection, mixing is slow; thus, the time evolution of the NO_x:HNO_3 ratio after convection depends largely on the partitioning of NO_x (between NO and NO_2), the concentration of OH, and the actinic flux.

Reactive nitrogen partitioning in the UT. The chemical sinks of UT NO_x are reaction with OH to produce HNO_3 (Eq. 1) and loss through N_2O_5 (Eqs. 2a and 2b), where M represents a third molecule (e.g., N_2, O_2) that absorbs the excess vibrational energy of the association reaction, followed by hydrolysis of N_2O_5 to produce HNO_3 (*36*). NO_x is regenerated by HNO_3 photolysis and reaction of OH with HNO_3 (and subsequent NO_3 photolysis to NO_2) (Eqs. 3 and 4).

$$NO_2 + OH + M \rightarrow HNO_3 + M \quad (1)$$

$$NO_2 + O_3 \rightarrow NO_3 + O_2 \quad (2a)$$

$$NO_3 + NO_2 + M \leftrightarrow N_2O_5 + M \quad (2b)$$

$$HNO_3 \xrightarrow{h\nu} OH + NO_2 \quad (3)$$

$$HNO_3 + OH \rightarrow NO_3 + H_2O \quad (4)$$

Assuming a diurnal steady state for HNO_3, the NO_x:HNO_3 ratio can be calculated as

$$\left(\frac{[NO_x]}{[HNO_3]}\right)_{\text{steady state}} = \frac{j_{HNO_3} + k_{HNO_3+OH}[OH]}{\left(k_{NO_2+OH}[OH] + 2k_{N_2O_5+H_2O}\frac{[N_2O_5]}{[NO_2]}\right) \times \left(\frac{[NO_2]}{[NO_3]}\right)} \quad (5)$$

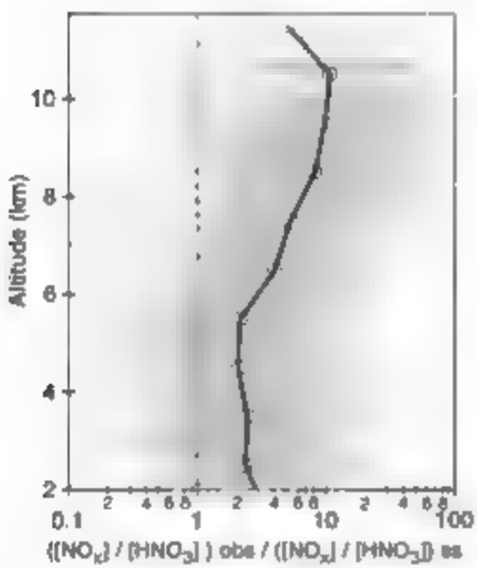


Fig. 2. Observed deviation of the NO_y:HNO_3 ratio from steady state as a function of altitude in the UT. The mean values within 1-km vertical bins are denoted by circles. The steady-state NO_y:HNO_3 ratio was calculated from measured NO_y, OH, and j_{HNO_3}. The grayscale data points were calculated from all observations taken during INTEX-NA.

where k is the reaction rate coefficient. NO_y:HNO_3 is expected to be larger than the steady-state value because wet scavenging removes HNO_3 faster than the time to reach steady state (*37*). Our observations show the NO_y:HNO_3 ratio to be much higher than the ratio described by Eq. 5 at altitudes greater than 6 km (Fig. 2). The difference between the observed ratio and that predicted by Eq. 5 grows with altitude, reaching a maximum at 10 km. Previous observations of NO_y and HNO_3 (either measured directly or calculated from observations of NO_y, peroxyacetylnitrate (PAN), and NO_y = NO_x + total peroxy nitrates + total alkyl nitrates + HNO_3 + 2 × N_2O_5 + other minor components)) have shown NO_y:HNO_3 to be significantly larger than the steady-state prediction in the UT (*30, 31, 38–42*). This has been shown to be primarily a result of convection and lightning reinitializing the system before steady state is achieved (*30, 31*). Although there are other hypotheses (*40–42*), we [like Jaegle *et al.* (*31*)] find no evidence for a mechanism other than convection responsible for holding NO_y:HNO_3 out of steady state in the UT.

Chemical signatures of convection. Figure 3 depicts one of many convectively influenced air masses sampled in the UT during INTEX-NA. Three distinct convective events (40 to 80 km wide) are identified by enhancements in NO_y:HNO_3 in Fig. 3A. Coincident enhancements are present in SO_2, an indicator of a recent boundary-layer source for this air, and in ultrafine condensation nuclei (UCN) ($3 < D_p < 10$ nm, where D_p is the particle diameter), an indicator of cloud detrainment (Fig. 3B) (*43, 44*). Sharp decreases in CO_2 also indicate the presence of boundary-layer air that has been depleted in CO_2 by photosynthetic activity (Fig. 3C) (*44*). Enhancements in CO, CH_3O, and various hydrocarbons relative to the surrounding UT air were also observed in these plumes, confirming that these parcels originate from the planetary boundary layer (PBL). Backward air trajectories, initialized along the flight track and mapped onto the spatial and temporal distribution of cloud-to-ground lightning strikes, indicate that this air mass was influenced by lightning about 1 day before DC-8 sampling (Fig. 3E). Such features with high NO_y:HNO_3 ratios were observed throughout the UT during INTEX-NA.

To assess the extent to which the UT over the eastern United States and Canada during the summer of 2004 was influenced by convection and to describe the chemical evolution of convective outflow, we used a constrained time-dependent photochemical box model to map the observed NO_y:HNO_3 ratio to the time since the ratio was last reinitialized (*45*). It was initialized with observations at 1-km vertical intervals from 6 to 12 km. The derived tuning indicator for the convectively influenced air sampled on 11 August 2004 is shown in Fig. 3D. The properties of the ensemble of our measurements are shown in Figs. 4 to 6.

The aerosol size distribution provides an independent indicator of air recently detrained from clouds. Cloud-processed air is depleted of aerosol surface area, permitting new particle formation in the outflow region (*43, 44*). Figure 4A depicts the fraction of condensation nuclei found in the 3- to 10-nm bin as a function of time since convective influence. The fraction of particles in this ultrafine mode is largest during the first few days, which confirms that the NO_y:HNO_3 ratio, and the tuning indicator derived from it, is reinitialized in the UT by cloud processing. Strong enhancements in CH_3OOH:H_2O_2, also an indicator of recent cloud processing (*33*), were observed during the first 2 days after cloud processing.

As expected, both elevated NO_x and suppressed HNO_3 are observed at short times (Fig. 4, B and C). Enhancement in NO_x during the first few days is indicative of convection of boundary-layer and/or lightning NO_x (*46*). The suppression of HNO_3 at short times is clear indication of HNO_3 scavenging during convection. Figure 4D confirms that reactive nitrogen (NO_y) is conserved during the chemical processing after convection, a fact that provides further support for the use of NO_y:HNO_3 as a marker representing time since convection.

Chemical processing in convective outflow. Mapping the ensemble of observations made throughout the UT onto the coordinate of time since convection allows us to assess the chemical and dynamical processes occurring after convection without attempting a Lagrangian convection study. In this analysis we concentrate on the time evolution of CO and O_3.

The time evolution of CO after detrainment into the UT is set by the abundance of OH and the rate at which the convective plume entrains air from the background UT (Fig. 5A). Because the chemical clock directly depends on HO_x, we constrained both OH and HO_2 to the observations as a function of NO_y and pressure in the time-dependent model used to generate time. As a result, we can iterate the model to determine the proper mixing rate of the convective plume by matching the modeled and observed time evolution of CO after convection. Using this approach for a series of long-lived species (e.g., CO, CH_4, CH_3OH), we calculated an average mixing rate of 0.05 – 0.02 day^{-1} after detrainment into the UT. This is in good agreement with the upper limit of 0.06 to 0.1 day^{-1} determined by Ray *et al.* from observations of convective plumes observed in the stratosphere (*47*). However, it is slower than the 2-day e-folding time scale determined by Wang *et al.* from observations in the UT (*32*). Because the DC-8 did not routinely sample in the turbulent environment directly surrounding convective outflow, this mixing rate likely reflects diffusive and shear-induced mixing subsequent to the initial turbulent mixing occurring during detrainment from the convective system.

Figure 5B shows the O_3 mixing ratio as a function of time since convection. We find that

on average convectively lofted air masses contain less O_3 than the background UT. This result is consistent with the observed vertical gradient in O_3 measured over the continental United States during INTEX-NA, with lower O_3 in the PBL than above (*48*). Rapid changes in the O_3 mixing ratio are observed during the first 2 days after detrainment, with the observed O_3 0 nmol mol^{-1} above the initial value by the end of day 2. The observed rate of increase slows exponentially with an asymptote at long times approaching zero and the O_3 mixing ratio approaching a constant value of 82 nmol mol^{-1}. This is a surprising result, as our model of the O_3 rate of change never approaches zero but continues to predict a net increase of 3 nmol mol^{-1} O_3 day^{-1} at the end of day 5 (*49*).

Constraints on the convective turnover rate of the UT. The convective turnover rate of the UT is critical for accurately describing NO_y, HO_x, and O_3 chemistry in the UT (*50*). However, at present there are few observation-based constraints available (either meteorological or chemical) to test the aggregate effects of convection in the current generation of global chemistry and climate models. To determine the convective turnover rate of the UT from the observations presented here, it is necessary to know with high confidence both the extent to which the UT is influenced by convection and the fraction of PBL air in the convectively influenced air masses.

To determine the fraction of PBL air contained in fresh convective outflow, we used observations of insoluble long-lived species. Assuming that we conducted a statistically unbiased sampling of both the boundary layer and free troposphere during INTEX-NA, we can calculate the fraction of PBL air present in fresh convection (f) according to

$$X_{UT(t=0)} = f\,X_{surface} + (1 - f)\,X_{UT} \quad (6)$$

where $[X]_{UT(t=0)}$ is the mean mixing ratio of species X in fresh convective outflow (as identified using our tuning indicator), $[X]_{UT}$ is the mean mixing ratio of species X in the UT (7.5 to 11.5 km), and $[X]_{surface}$ is the mean mixing ratio of species X between 0 and 1.5 km. Using observations of CO, CO_2, CH_3OH, CH_4, and C_3H_8, we calculated the fraction of PBL air in fresh convection to be 0.19 ± 0.05, 0.11 ± 0.03, 0.26 ± 0.05, 0.15 ± 0.05, and 0.34 ± 0.09, respectively. We calculated a weighted average for the fraction of PBL air in convective outflow of 0.17 ± 0.02 by weighting each value by the inverse square of its uncertainty. This result implies that convectively lofted PBL air rapidly entrains the surrounding air either during ascent or in the turbulent environment of the detraining flow. These results are consistent with the observations of Ray *et al.*, who determined the fraction of tropospheric air in convective plumes sampled in the stratosphere to be between 0.1 and 0.4 (*47*). However, our results suggest a smaller fraction than the observations of Cohan *et al.*, who calculated the fraction of boundary-layer air in fresh convective outflow to be between 0.32 and 0.64 (*33*), and the modeling studies of Mahowald *et al.*, who calculated the fraction of PBL air present in the convective outflow region of a supercell storm 10 hours after storm initialization to be 0.26 (*51*).

Figure 6A shows the normalized frequency distribution of the time since convection, based on the NO_y/HNO_3 ratio. We found that 54% of the air between 7.5 and 11.5 km was influenced by convection during the past 2 days. The convective outflow was strongest between 9.5 and 10.5 km, where the fraction of air that is less than 2 days old exceeds 69%. The vertical dis-

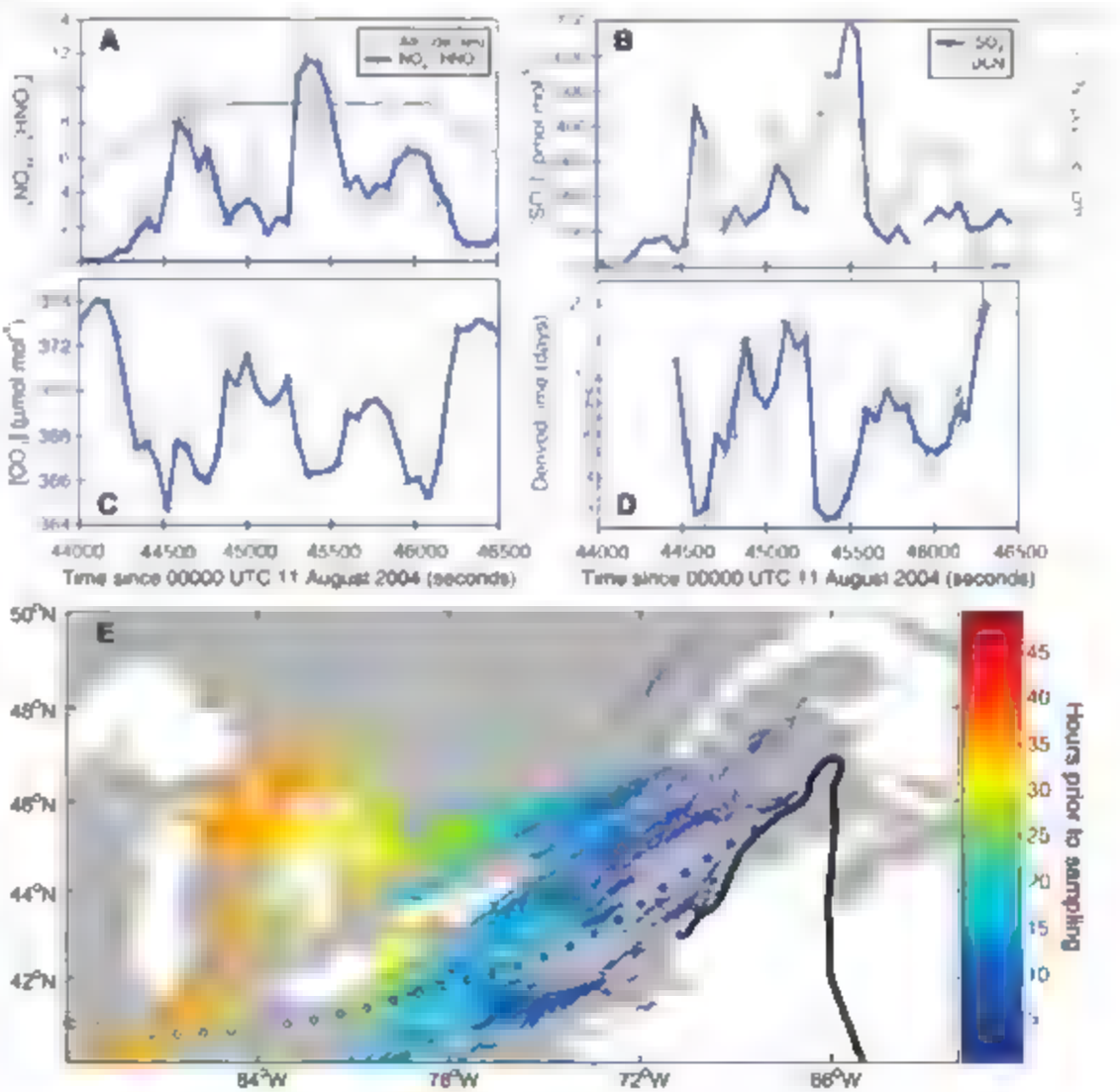


Fig. 3. (**A** to **D**) Time series of measurements taken in the vicinity of recent convective activity on 11 August 2004 between 5 and 9 km. (**A**) Three distinct convective plumes, each indicated by a sharp increase in the NO_y/HNO_3 ratio. (**B** and **C**) Coincident enhancements in SO_2 and UCN (3 nm > D_p > 10 nm) and coincident sharp drops in CO_2, indicative of the convective lofting of boundary-layer air depleted in CO_2. The derived time since the sampled air mass had been influenced by convection is shown in (**D**). (**E**) NLDN lightning hits (small dots) on 10 and 11 August. The color scale represents the time of the hit (hours) before aircraft sampling. The DC-8 sampling location corresponding to measurements shown in the top panel is located on the Maine–New Brunswick border (46°N, 67°W). The 2-day back-trajectory [initialized at the point of the second convective plume shown in (**A**)] is also color-coded by time before DC-8 sampling (circles with black edges). The DC-8 flight track on 11 August 2004 is shown by the heavy black line.

tribution presented here is consistent with previous analyses of convective outflow to the UT from individual storms (*4*, *52*) and the vertical distribution of convectively influenced laminae observed in O_3 sonde data from the summer of 2004 over the northeastern United States. The shift toward longer times between 10.5 and 11.5 km suggests two possibilities: that convective cloud tops on average do not extend higher than 10.5 km over the mid-latitudes during summer (*53*), or that transport of stratospheric air rich in HNO_3 contributes to keeping the NO_y:HNO_3 ratio low at altitudes higher than 10.5 km.

To constrain the turnover rate of the UT from the ensemble statistics generated from our calculated time since convective influence (Fig. 6A), we constructed a two-dimensional (2D) model of the UT. On the basis of typical wind speeds, we assume that it takes 4 days for any individual model point to pass through the sampling region and that each point has not been influenced by convection upon entering the sampling window. Every 6 hours, we represented convection by randomly reinitializing the age of x% of the points to 0 [the value of x is determined by the turnover rate (varied between 0.05 and 0.2 day^{-1}) and the fraction of PBL air contained in fresh convection (assumed to be a constant at 0.17)], and we then diluted each point with the mean value of the adjacent 8 points at the rate of 0.05 day^{-1}.

Figure 6B depicts the observed and three calculated normalized frequency distributions of time since convective influence between 7.5 and 11.5 km. The shape of the distribution suggests that UT air sampled during INTEX-NA was strongly influenced by convection, and that convectively lofted plumes did not have sufficient time to either mix or age before sampling; instead, air was transported to the east out of the domain. Calculated frequency distributions of time since convection, obtained by collecting the points in the eastern half of the 2D UT model analysis (where we sampled most frequently), are also shown in Fig. 6B. Assuming the DC-8 made a statistically unbiased sampling of the continental UT during summer, the best match among the three model calculations and observations would imply a convective turnover rate between 0.1 and 0.2 day^{-1}. However, if we assume that the DC-8 had a positive bias toward sampling fresh convection, our observed frequency distributions are most consistent with a convective turnover rate closer to 0.1 day^{-1} (*54*).

For comparison, the Goddard Earth Observing System (GEOS-4) data assimilation model detrainment cloud mass flux between 400 and 200 hPa (~7.2 to 11.8 km) for the domain (80° to 100° W for 30° to 35° N and 70° to 100° W for 35° to 50°N) between 1 July and 15 August 2004 was 0.0085 kg m^{-2} s^{-1}. This corresponds to a turnover rate of 0.37 day^{-1} (using a column mass of 1.9×10^3 kg m^{-2} between 7.2 and 11.8 km). Further investigation is needed to understand the source of the difference between our observed turnover rate and the one derived from the model.

Conclusions. Our observations provide unique constraints on (i) the extent to which convection perturbs the continental UT during summer, (ii) the fraction of boundary layer air present at convective outflow, and (iii) the convective overturning rate of the UT. In addition, the chemical clock described here defines a coordinate that can be used to assess the chemistry occurring downwind of convective injection. These direct measures of atmospheric rates present a new opportunity for quantitative tests of model representations of processes governing UT ozone, convection, and lightning and their impact on climate.

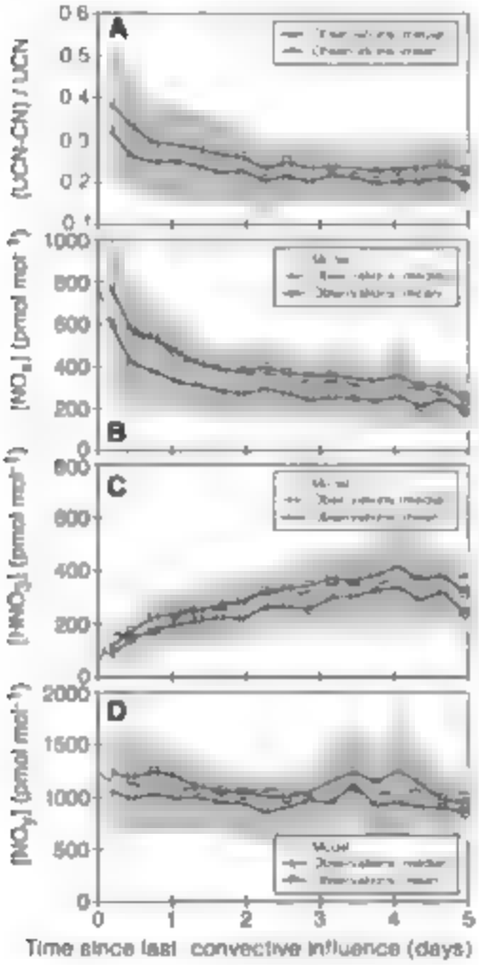


Fig. 4. Observations of (**A**) the fraction of ultrafine condensation nuclei [number density of aerosol (3 to 10 nm)/total aerosol number density], (**B**) NO_x, (**C**) HNO_3, and (**D**) NO_y as a function of modeled time since convective influence. Means and medians of the observations, within 8-hour bins, are shown along with the interquartile range (shaded region). Results from the time-dependent box model, initialized at 10 km and at noon, are shown with dashed lines for the gas-phase species [(B) to (D)].

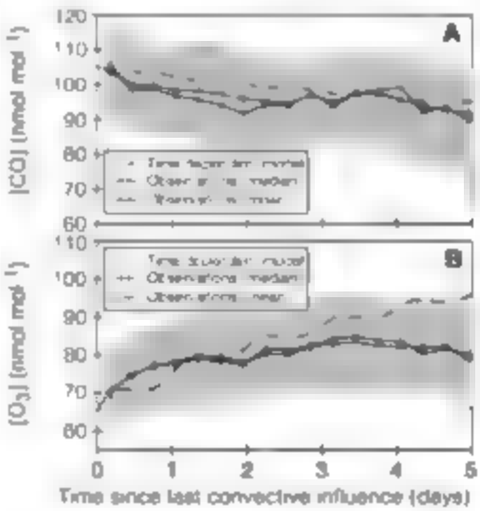


Fig. 5. Observations of CO (**A**) and ozone (**B**) as a function of modeled time since cloud processing in the UT. Means and medians of the observations, within 8-hour bins, are shown along with the interquartile range (shaded region). Results from the time-dependent box model, initialized at 10 km and at noon, are shown with dashed lines.

References and Notes

1. R. B. Chatfield, P. J. Crutzen, *J. Geophys. Res.* **89**, 7111 (1984).
2. R. R. Dickerson *et al.*, *Science* **235**, 460 (1987).
3. K. E. Pickering, R. R. Dickerson, G. J. Huffman, J. F. Boatman, A. Schanot, *J. Geophys. Res.* **93**, 759 (1988).
4. A. M. Thompson *et al.*, *J. Geophys. Res.* **99**, 18703 (1994).
5. W. J. Collins, R. G. Derwent, C. E. Johnson, D. S. Stevenson, *Q. J. R. Meteorol. Soc.* **128**, 991 (2002).
6. J. E. Dye *et al.*, *J. Geophys. Res.* **105**, 10023 (2000).
7. S. A. Rutledge, R. A. Houze, M. I. Biggerstaff, T. Matejka, *Mon. Weather Rev.* **116**, 1409 (1988).
8. H. Huntrieser, H. Schlager, C. Feigl, H. Holler, *J. Geophys. Res.* **103**, 28247 (1998).
9. B. Ridley *et al.*, *J. Geophys. Res.* **109**, D17305 (2004).
10. *Scientific Assessment of Ozone Depletion* (World Meteorological Organization, Geneva, 1995).
11. L. Jaeglé, L. Steinberger, R. V. Martin, K. Chance, *Faraday Discuss.* **130**, 407 (2005).
12. A. J. DeCaria, K. E. Pickering, G. L. Stenchikov, L. E. Ott, *J. Geophys. Res.* **110**, D14303 (2005).
13. J. Lelieveld, P. J. Crutzen, *Science* **264**, 1759 (1994).
14. M. G. Lawrence, R. von Kuhlmann, M. Salzmann, P. J. Rasch, *Geophys. Res. Lett.* **30**, 10.1029/2003GL017644 (2003).
15. M. Gauss *et al.*, *J. Geophys. Res.* **108**, 10.1029/2002JD002624 (2003).
16. K. E. Pickering *et al.*, *J. Geophys. Res.* **95**, 14049 (1990).
17. I. Folkins, K. K. Kelly, E. M. Weinstock, *J. Geophys. Res.* **107**, 10.1029/2002JD002185 (2002).
18. M. T. Chahine, *Nature* **359**, 373 (1992).
19. J. A. Thornton, P. J. Wooldridge, R. C. Cohen, *Anal. Chem.* **72**, 528 (2000).
20. P. A. Cleary, P. J. Wooldridge, R. C. Cohen, *Appl. Opt.* **41**, 6950 (2002).
21. R. W. Talbot *et al.*, *Geophys. Res. Lett.* **26**, 3057 (1999).
22. J. D. Crounse, K. A. McKinney, A. J. Kwan, P. O. Wennberg, *Anal. Chem.* **78**, 6726 (2006).
23. I. C. Faloona *et al.*, *J. Atmos. Chem.* **47**, 139 (2004).
24. A. D. Clarke *et al.*, *J. Geophys. Res.* **109**, D15S09 (2004).
25. R. E. Shetter, M. Muller, *J. Geophys. Res.* **104**, 5647 (1999).
26. G. W. Sachse, G. F. Hill, L. O. Wade, M. G. Perry, *J. Geophys. Res.* **92**, 2071 (1987).
27. S. A. Vay *et al.*, *J. Geophys. Res.* **104**, 5663 (1999).
28. H. B. Singh, W. H. Brune, J. H. Crawford, D. J. Jacob, P. B. Russell, *J. Geophys. Res.* **111**, D24S01 (2006).
29. R. Sander, *Compilation of Henry's Law Constants for Inorganic and Organic Species of Potential Importance in Environmental Chemistry (Version 3)* (1999), available at www.henrys-law.org.
30. M. J. Prather, D. J. Jacob, *Geophys. Res. Lett.* **24**, 3189 (1997).

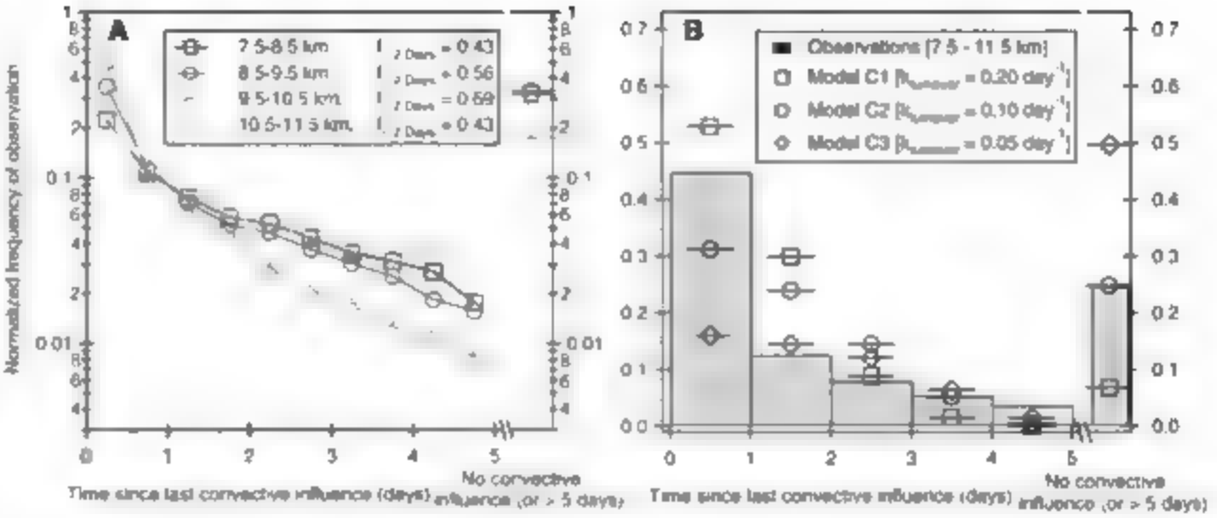


Fig. 6. (**A**) Normalized frequency distribution in the time since convective influence, as calculated from observations of the NO_y/HNO_3 ratio made during summer 2004. Calculations are separated into 1-km altitude bins (range 7.5 to 11.5 km). The fraction of air that had been influenced by convection within the past 2 days ($f_{<2\ days}$) is shown. (**B**) Comparison of observed frequency distribution (7.5 to 11.5 km) with various modeled representations of the convective turnover rate.

31. L. Jaeglé *et al. Geophys. Res. Lett.* **25**, 1705 (1998).
32. Y. Wang *et al., Geophys. Res. Lett.* **27**, 369 (2000).
33. D. S. Cohan, M. G. Schultz, D. J. Jacob, B. G. Heikes, D. R. Blake, *J. Geophys. Res.* **104**, 5717 (1999).
34. S. Smyth *et al., J. Geophys. Res.* **101**, 1743 (1996).
35. S. A. McKeen, S. C. Liu, *Geophys. Res. Lett.* **20**, 2363 (1993).
36. F. J. Dentener, P. J. Crutzen, *J. Geophys. Res.* **98**, 7149 (1993).
37. F. Giorgi, W. L. Chameides, *J. Geophys. Res.* **90**, 7872 (1985).
38. D. D. Davis *et al., J. Geophys. Res.* **101**, 2111 (1996).
39. D. J. Jacob *et al., J. Geophys. Res.* **101**, 24235 (1996).
40. R. B. Chatfield, *Geophys. Res. Lett.* **21**, 2705 (1994).
41. A. Fabazadeh *et al., Geophys. Res. Lett.* **25**, 4185 (1998).
42. D. A. Hauglustaine, B. A. Ridley, S. Solomon, P. G. Hess, S. Madronich, *Geophys. Res. Lett.* **23**, 2609 (1996).
43. A. D. Clarke *et al., J. Geophys. Res.* **104**, 5735 (1999).
44. H. Huntrieser *et al., J. Geophys. Res.* **107**, 10.1029/2000JD000209 (2002).
45. See supporting material on *Science* Online.
46. O. R. Cooper *et al., J. Geophys. Res.* **111**, D24S05 (2006).
47. E. A. Ray *et al., J. Geophys. Res.* **109**, D18304 (2004).
48. T. Haul, P. Schulte, R. Ahrett, H. Schlager, *J. Geophys. Res.* **100**, 22957 (1995).
49. Net ozone concentration change of 0 nmol mol^{-1} day^{-1} could be achieved if the air parcel (i) subsided to where H_2O abundances were large enough to provide a sink of O through O^1D hat balanced production from NO + HO_2 (~6 km), (ii) entrained air containing lower O_3 mixing ratios, or (iii) contained additional O_3 loss terms beyond NO_y, HO_x, and H_2O (via O^1D removal). To match the deviation between the model and measurement, we would require an additional 2 to 3 nmol mol^{-1} day^{-1} of chemical ozone loss. In order for mixing to explain the deviation, air of lower O_3 would need to be mixed into the air parcel. As shown in Fig. 5B, the only air in the UT containing significantly less O_3 is that which is pumped directly from the PBL. Although mixing fresh and aged outflow could help to explain the discrepancy in O_3, it is inconsistent with the observed decay in CO at long times (2 to 5 days).
50. D. Rind, J. Lerner, *J. Geophys. Res.* **101**, 12667 (1996).
51. G. L. Mullendore, D. R. Durran, J. R. Holton, *J. Geophys. Res.* **110**, D06113 (2005).
52. K. E. Pickering, Y. S. Wang, W. K. Tao, C. Price, J. F. Muller, *J. Geophys. Res.* **103**, 31203 (1998).
53. W. B. Rossow, http://isccp.giss.nasa.gov/products/isccpDsets.html (2006).
54. We used 10-day back-trajectories to National Weather Service Global Forecast System (GFS)–derived convection measurements and NLDN-measured lightning strikes to assess the fraction of time that the DC-8 sampled either convection- or lightning-influenced air. Using the GFS statistics, we calculated that 63% of the sampled air on INTEX-NA had encountered convection and ~57% had been influenced by lightning during the past 2 days. When considering the entire INTEX-NA sampling domain (both in space and time), convection was present in 12.5% of the grid points. This is substantially smaller than the percentage of observations within 6 hours of convection (21.4%), which suggests that the DC-8 had a positive bias toward sampling fresh convection. This bias is reflected in the sharp drop in population between day 1 and day 2 (Fig. 6A). Correcting for this bias had little effect on our assessment of the fraction of air less than 2 days old, lowering our results from 0.43, 0.56, 0.69, and 0.43 to 0.38, 0.50, 0.62, and 0.39 at 8, 9, 10, and 11 km, respectively.
55. We thank the flight and ground crews of the NASA DC-8 aircraft and the entire INTEX-NA science team for their contributions during the 2004 intensive field campaign. A. M. Thompson, I. Folkins, M. G. Lawrence, and D. Allen for helpful discussions, B. Kucsera for help with the GEOS-4 calculations, and W. H. Brune and X. Ren for OH and HO_2 data. NLDN data were collected by Vaisala Thunderstorm and provided to the INTEX Science Team by the Global Hydrology Resource Center at NASA Marshall Space Flight Center. Work at UC Berkeley was supported by NASA grants NNG05GH196 and NAG5 13668. The INTEX-NA field program was supported by the NASA-ESE Tropospheric Chemistry Program.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1134548/DC1
Materials and Methods
Figs. S1 to S8
Table S1
References

31 August 2006; accepted 19 December 2006
Published online 4 January 2007
10.1126/science.1134548
Include this information when citing this paper.

A Membrane Receptor for Retinol Binding Protein Mediates Cellular Uptake of Vitamin A

Riki Kawaguchi,[1,2] Jiamei Yu,[1] Jane Honda,[1] Jane Hu,[2] Julian Whitelegge,[1,4] Peipei Ping,[1,5] Patrick Wiita,[1] Dean Bok,[2,4,6] Hui Sun[1,2,4]*

Vitamin A has diverse biological functions. It is transported in the blood as a complex with retinol binding protein (RBP), but the molecular mechanism by which vitamin A is absorbed by cells from the vitamin A–RBP complex is not clearly understood. We identified in bovine retinal pigment epithelium cells STRA6, a multitransmembrane domain protein, as a specific membrane receptor for RBP. STRA6 binds to RBP with high affinity and has robust vitamin A uptake activity from the vitamin A–RBP complex. It is widely expressed in embryonic development and in adult organ systems. The RBP receptor represents a major physiological mediator of cellular vitamin A uptake.

Vitamin A and its derivatives are essential for vision (*1*) and many other biological processes, because they are involved in the proliferation and differentiation of many cell types throughout life (*2*, *3*). The majority of dietary vitamin A is stored in the liver. The principal physiological carrier of vitamin A (retinol) in the blood for delivery to other organs is retinol binding protein (RBP) (*4*). RBP is responsible for a well-regulated transport system that provides an evolutionary advantage by helping vertebrates adapt to fluctuations in vitamin A in natural environments (*5*). RBP also functions as a signal in insulin resistance (*6*). Loss of RBP makes mice extremely sensitive to vitamin A deficiency, because the hepatic vitamin A store can no longer be mobilized (*7*). Even with a nutritionally complete diet, RBP knockout mice have dramatically lower serum vitamin A concentrations, similar to the concentrations in the

later stages of vitamin A deficiency in humans. Given the role of vitamin A in immune regulation and the susceptibility of vitamin A deficient children to infection before visual symptoms (*8*), it is likely that the immune system is compromised in RBP knockout mice even under vitamin A sufficient conditions. Indeed, the circulating immunoglobulin amount in these mice is half of that in wild-type mice, even under vitamin A sufficiency (*9*). In humans, loss of RBP function causes a dark adaptation defect and progressive atrophy of the retinal pigment epithelium (RPE) at young ages (*10*). Under conditions of vitamin A deficiency at which wild-type mice behave normally, RBP knockout mice have severe developmental defects in embryos (*11*) and rapid vision loss in adults after merely a week of vitamin A deficiency (*7*). In addition, these mice rapidly develop testicular defects (*12*).

It was first shown in the 1970s that there exists a specific cell surface receptor for RBP on the RPE and on intestinal epithelial cells (*13*–*16*). During the past 3 decades, evidence has also accumulated for the existence of the RBP receptor on other tissue or cell types, including the placenta (*17*–*20*), choroid plexus (*15*, *21*), testis (*18*, *22*), and macrophages (*23*). The cell surface RBP receptor not only specifically binds to RBP but also mediates vitamin A uptake from vitamin A–loaded RBP (holo-RBP) (*15*, *17*, *20*, *22*, *23*). Here, we report the identification of the RBP receptor as STRA6, a widely expressed multitransmembrane domain protein. STRA6 met all three criteria expected of the RBP receptor. First, it conferred RBP binding to transfected cells. Second, it mediated cellular uptake of vitamin A. Third, it was localized to the cellular locations expected of the RBP receptor in native tissues.

Identification of the RBP receptor as STRA6. Potential obstacles to purifying the RBP receptor included the fragility of the receptor protein and the transient nature of binding. We designed a strategy to stabilize the RBP-receptor interaction, and this strategy permitted high affinity purification of the RBP-receptor complex. The strategy combined the 6XHis-tag (His tag) nickel system and a bifunctional cross-linker with an amine-reactive group and a photoreactive group (Fig. 1A). We used bovine RPE cells as the starting material because they are known to express the RBP receptor at high concentrations (*13*). Because N-terminal alkaline phosphatase (AP)–tagged RBP was able to specifically bind to its receptor on RPE cells with a staining pattern expected for basolateral membranes (Fig. 1B), we created His-tagged RBP (His-RBP) by tagging at the N terminus of RBP. After binding purified His-RBP protein conjugated with the cross-linker to RPE membranes, ultraviolet (UV) cross-linking, membrane solubilization, and nickel resin purification, we observed a covalent protein complex of about 80 kD containing RBP and its putative receptor (Fig. 1C). Mass spectrometry analysis revealed that the RBP binding protein in the complex was STRA6, a protein that has 11 putative transmembrane domains but has no homology to any protein with known function (fig. S1). *STRA6* is a retinoic acid–induced gene in P19 embryonic carcinoma cells (*24*).

STRA6 mediates RBP binding and vitamin A uptake from holo-RBP. To test whether STRA6 could confer RBP binding to transfected cells, we transfected bovine STRA6 cDNA into COS-1 cells. Bovine STRA6–transfected but not untransfected cells bound to AP-RBP with high affinity (K_d = 59 nM) (Fig. 2, A and B). An excess of unlabeled RBP blocked this binding activity (Fig. 2B). To test whether STRA6 could enhance cellular vitamin A uptake from holo-RBP, we performed both ^{3}H-retinol–based assays (Fig. 2C) and high-performance liquid chromatography (HPLC)–based assays (Fig. 2, D to H). STRA6 expression dramatically increased cellular vitamin A uptake from holo-RBP. Vitamin A is preferentially stored as retinyl esters after uptake from holo-RBP (*25*). Lecithin retinol acyltransferase (LRAT) is the enzyme that converts retinol into retinyl ester (*26*) and plays an important role in vitamin A storage (*27*, *28*). Thus, we included LRAT expression in most vitamin A uptake assays. By using an assay based on ^{3}H-retinol-RBP, we found that STRA6 ac-

[1]Department of Physiology, David Geffen School of Medicine at UCLA, 650 Charles E. Young Drive South, Los Angeles, CA 90095, USA. [2]Jules Stein Eye Institute, David Geffen School of Medicine at UCLA, 100 Stein Plaza, 650 Charles E. Young Drive South, Los Angeles, CA 90095, USA. [3]Pasarow Mass Spectrometry Laboratory, Semel Institute, David Geffen School of Medicine at UCLA, 760 Westwood Plaza, Los Angeles, CA 90024, USA. [4]Brain Research Institute, UCLA, Los Angeles, CA 90095, USA. [5]Department of Medicine, David Geffen School of Medicine at UCLA, 650 Charles E. Young Drive South, Los Angeles, CA 90095, USA. [6]Department of Neurobiology, David Geffen School of Medicine at UCLA, 650 Charles E. Young Drive South, Los Angeles, CA 90095, USA

*To whom correspondence should be addressed. E-mail: hsun@mednet.ucla.edu

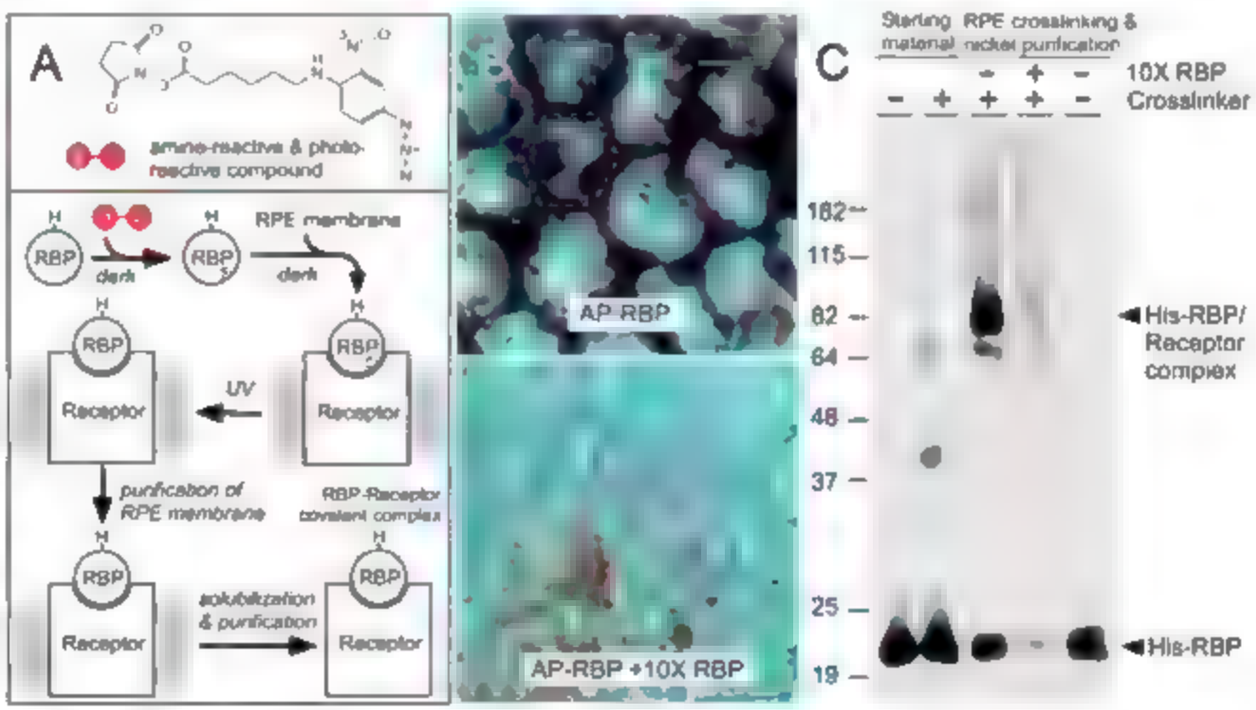


Fig. 1. (**A**) Schematic diagram of the strategy used to identify the RBP receptor. The structure of the amino-reactive and photo-reactive cross-linker is shown on the top and illustrated as red symbols in the diagram. The letter H above the symbol for RBP represents the 6XHis tag at the N terminus. Shaded rectangles represent membrane. (**B**) Binding of AP-RBP to freshly dissected bovine RPE cells (top). AP reaction product is deep purple. Presence of 10-fold excess RBP abolished the labeling (bottom). The hexagonal shape of an RPE cell is outlined in red dotted lines in each picture. Scale bars indicate 10 μm. (**C**) Immunoblot analysis using an antibody against RBP. Molecular weight marker is shown on the left (in kD). The His-RBP monomer band represents His-RBP that was bound to RPE membrane but not cross-linked to the receptor. Untagged RBP competition in lane 4 prevented His-RBP from binding to RPE membrane and therefore lowered signals for both the complex and the monomer band.

tivity was most evident in the presence of LRAT (Fig. 2C), because LRAT makes it possible to store vitamin A at high concentration in the form of retinyl esters, the more stable derivative of vitamin A. In these assays, nearly all specific vitamin A uptake activity, which was sensitive to blocking by an excess of unlabeled RBP, was due to STRA6. Vitamin A uptake by COS-1 cells transfected with STRA6 and LRAT was highly efficient. When the concentration of ^{3}H-retinol-RBP was 3 nM in the medium and cell density was 1.8×10^6 cells/mL, about 25% of the total radioactivity was taken up by cells in 1 hour, and about 50% of the total radioactivity was taken up by cells in 2 hours. The amount of retinol taken up by cells in 1 hour is about 4.6 times the moles of RBP bound to the cell surface at steady state. We also used HPLC to analyze vitamin A uptake mediated by STRA6 (Fig. 2, D to H). In this assay, untransfected COS-1 cells accumulated little retinyl ester. COS-1 cells transfected with LRAT alone did accumulate a small amount of retinyl esters, but cells cotransfected with STRA6 and LRAT showed a 15-fold higher amount of retinyl ester accumulation than that of cells transfected with LRAT alone (Fig. 2, D to F). At the saturating retinol-RBP concentration (0.44 μM) used for the HPLC assays, the amount of retinol taken up by cells in 1 hour is about 14 times the moles of RBP bound to the cell surface at steady state. Holo-RBP in the blood is in complex with the thyroxine binding protein, transthyretin (TTR). Although TTR blocks the vitamin A exit site in the holo-RBP-TTR complex (29), the fact that tissues take up vitamin A from holo-RBP-TTR in the blood in vivo suggests that TTR cannot completely inhibit tissue vitamin A uptake from holo-RBP. When we assayed vitamin A uptake by using human serum as the source of the holo-RBP-TTR complex, STRA6 could still efficiently take up vitamin A from this complex (Fig. 2, G and H).

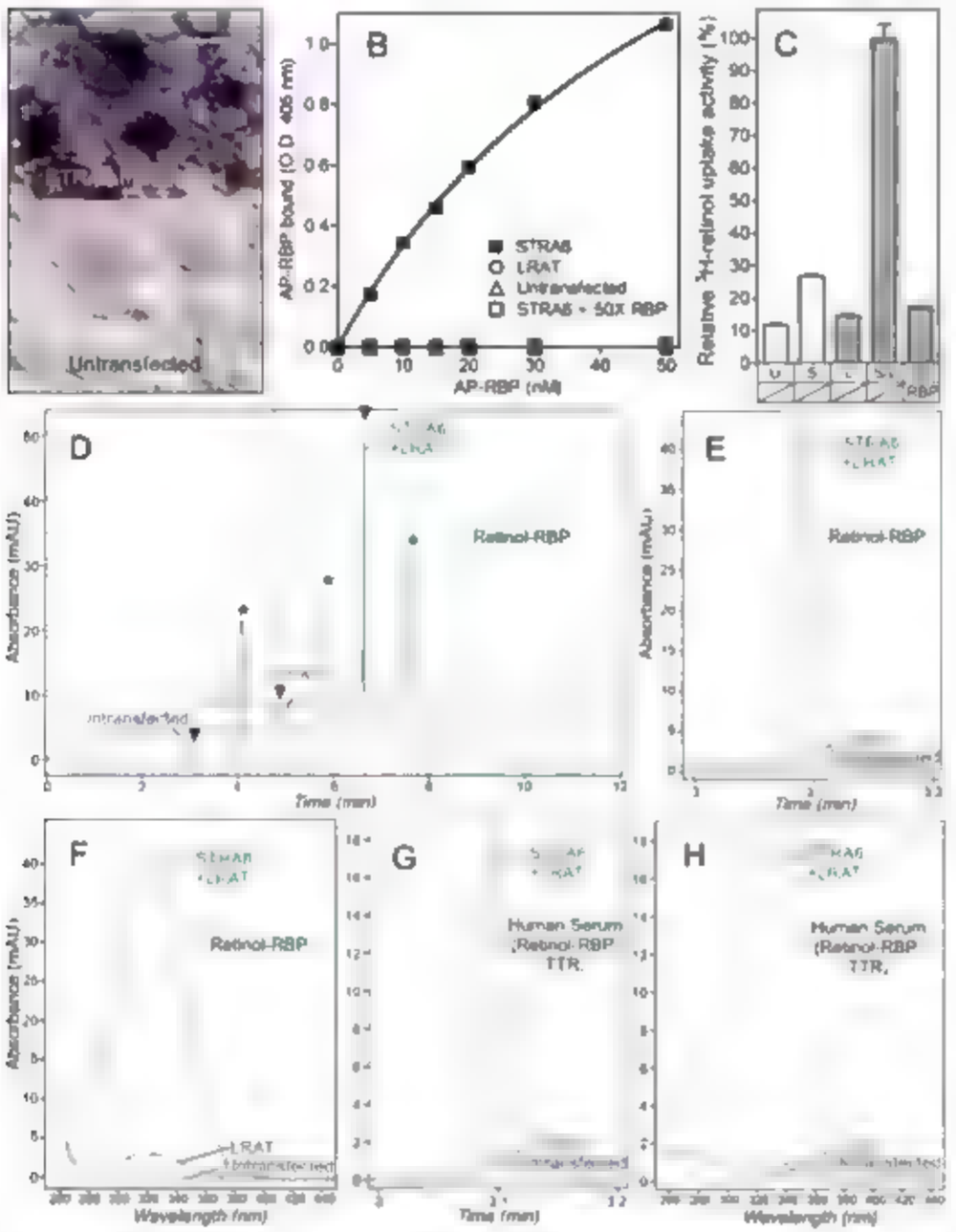


Fig. 2. RBP binding and vitamin A uptake activity of STRA6. (**A**) Binding of RBP to live COS-1 cells transfected with STRA6 (top) but not to untransfected cells (bottom). The AP reaction product is deep purple. Scale bar, 20 μm. (**B**) Concentration dependence of the binding of AP-RBP to live COS-1 cells transfected with STRA6. Controls are LRAT-transfected cells, untransfected cells, and STRA6-transfected cells with 50× unlabeled RBP during binding. The plots for these three controls are superimposed on this graph. (**C**) Comparison of vitamin A uptake activities from ^{3}H-retinol/RBP. U, S, L, and S/L denote untransfected cells and cells transfected with STRA6, LRAT, and both, respectively. RBP, 100× unlabeled holo-RBP. The activity of S/L cells is defined as 100%. Error bars indicate SD. (**D** to **H**) HPLC assays of vitamin A uptake for COS-1 cells transfected with STRA6 and LRAT (green) or LRAT alone (red) and for untransfected cells (blue). AU, absorbance units. (**D**) HPLC profile of hexane extracts of COS-1 cells after vitamin A uptake from holo-RBP. Arrowheads indicate peaks representing retinyl ester. Peaks representing hexane-soluble peptides (asterisks) serve as an internal control. (**E**) Overlay of the retinyl ester peaks in (D). (**F**) Overlay of absorption spectra of the retinyl ester peaks in (D). (**G**) Overlay of the retinyl ester peaks from similar HPLC runs as shown in (D) but using 25% human serum as the source of holo-RBP. (**H**) Overlay of absorption spectra of the peaks in (G).

Because retinoic acid up-regulates STRA6 in certain cancer cell lines such as WiDr human colon adenocarcinoma cells (*24*, *30*), we used WiDr cells as an independent model to study vitamin A uptake mediated by STRA6. We increased STRA6 expression in WiDr cells by retinoic acid treatment and decreased its expression by RNA interference (RNAi). Consistent with STRA6's function in mediating vitamin A uptake from holo-RBP, its increased expression by retinoic acid treatment enhanced vitamin A uptake activity, whereas its decreased expression by specific RNAi knockdown suppressed vitamin A uptake activity (Fig. 3A). In addition, we performed RNAi knockdown experiments on primary bovine RPE cultures. Again, a decrease in STRA6 expression suppressed RPE's vitamin A uptake activity from holo-RBP (Fig. 3B).

Because STRA6 is not homologous to any protein of known function and has never been characterized at the structural level, we used an unbiased strategy to test the effects of mutations on STRA6 function. We produced and characterized 50 random missense mutants of STRA6 and found that 3 of them have substantial loss of vitamin A uptake activity (Fig. 3, C to E). The locations of these mutations in the putative transmembrane topology model of STRA6 are shown in fig. S1B. Mutants 2 and 3 (M2 and M3) were not expressed at the cell surface, which may be sufficient to account for the dramatic loss of RBP binding and vitamin A uptake activity from holo-RBP. Mutant 1 (M1) was normally expressed at the cell surface but had substantially reduced RBP binding and vitamin A uptake activity. The loss of RBP binding for M1 may account for its lower vitamin A uptake activity.

Characterization of STRA6-mediated vitamin A uptake from holo-RBP. We found that STRA6-mediated vitamin A uptake is specific to RBP. STRA6 could not enhance cellular uptake of vitamin A when vitamin A was bound to bovine serum albumin (BSA) or β-lactoglobulin (Fig. 4A). STRA6-mediated vitamin A uptake was saturable with regard to retinol-RBP concentration (Fig. 4B and fig. S2). A variety of studies from the past 3 decades had indicated that endocytosis is not required for RBP receptor–mediated vitamin A uptake from holo-RBP (*14*, *17*, *20*, *22*–*25*). In agreement with these studies, metabolic inhibitors that effectively inhibited endocytosis did not inhibit STRA6-mediated vitamin A uptake by STRA6-transfected COS-1 cells or retinoic acid stimulated WiDr cells (Fig. 4C and 4D). Membranes prepared from transfected cells have vitamin A uptake activity similar to that of live cells (Fig. 4E). The robust activity of this cell-free system also argues against an endocytosis-based mechanism. We also designed an assay that monitored both RBP and vitamin A that remains bound to RBP in the

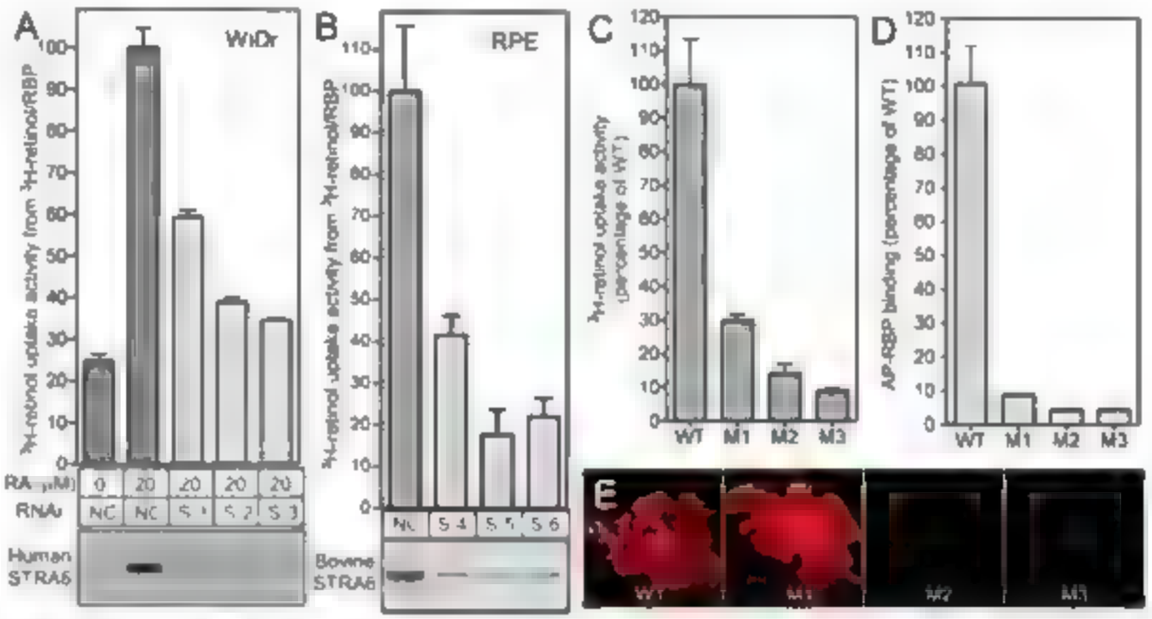


Fig. 3. STRA6 RNAi in native cell types and STRA6 mutagenesis. (**A**) Retinoic acid (RA) treatment of WiDr cells increases the expression of STRA6 and the vitamin A uptake activity of WiDr cells, whereas STRA6 RNAi treatment suppresses STRA6 expression and vitamin A uptake activity. NC is negative control RNAi oligonucleotide (oligo). S-1, S-2, and S-3 are RNAi oligos for human STRA6. Error bars indicate SD. (**B**) STRA6 RNAi treatment suppresses STRA6 expression and vitamin A uptake activity of primary bovine RPE culture. S-4, S-5, and S-6 are RNAi oligos for bovine STRA6. (**C**) Comparison of vitamin A uptake activities of COS-1 cells transfected with bovine STRA6 wild-type (WT) or mutants Gln^{310}→His^{310}/Leu^{315}→Pro^{315} (M1), Leu^{256}→His^{256}/Tyr^{336}→His^{336} (M2), and Leu^{297}→Arg^{297} (M3). Activity of cells transfected with WT STRA6 is defined as 100%. (**D**) Comparison of AP-RBP binding activities of COS-1 cells transfected with bovine STRA6 WT or mutant M1, M2, or M3. Activity of cells transfected with WT STRA6 is defined as 100%. (**E**) Surface expression of wild-type or mutant STRA6 as assayed by immunostaining of transfected live COS-1 cells. Scale bar, 10 μm.

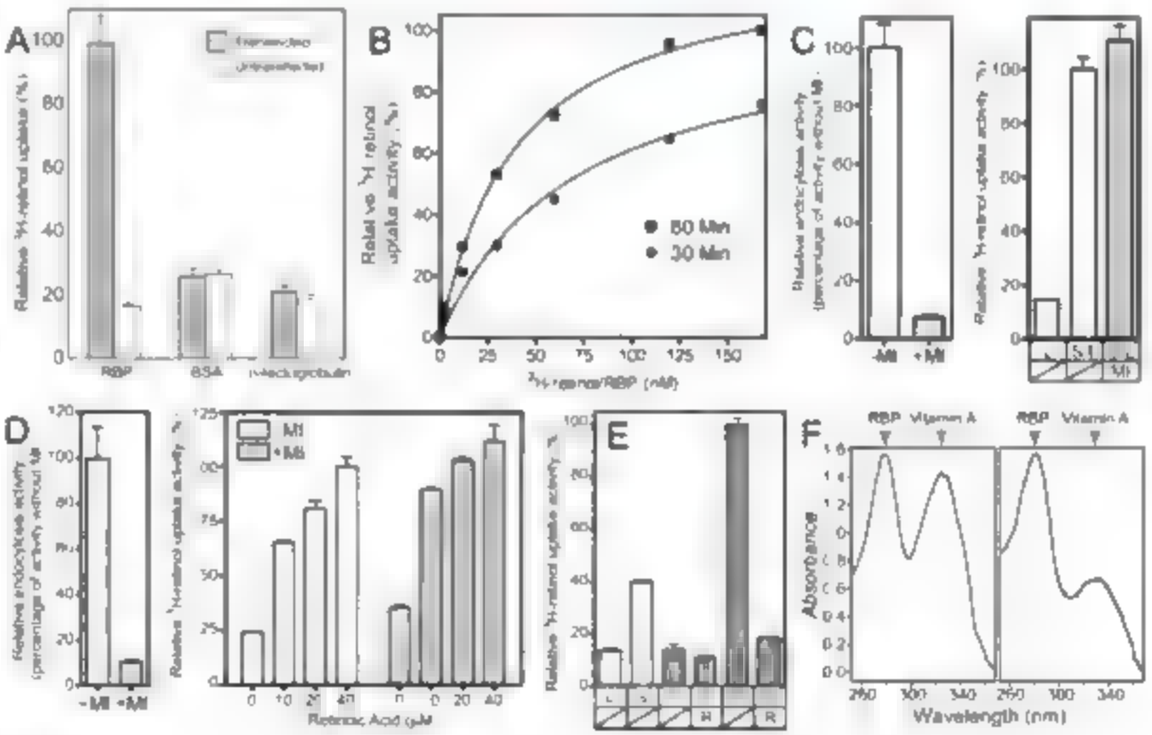


Fig. 4. Characterization of STRA6-mediated vitamin A uptake from holo-RBP. (**A**) STRA6 mediates ^{3}H-retinol uptake from ^{3}H-retinol bound to RBP, but not ^{3}H-retinol bound to BSA or β-lactoglobulin. "Transfected" indicates transfection with STRA6 and LRAT. Retinol uptake activity of transfected cells from holo-RBP is defined as 100%. Error bars indicate SD. (**B**) STRA6-mediated vitamin A uptake is saturable with regard to retinol-RBP concentration. The highest retinol uptake activity at 60 min is defined as 100%. (**C**) MI (a mixture of metabolic inhibitors 2-deoxyglucose and sodium azide) effectively inhibits horseradish peroxidase (HRP) endocytosis of COS-1 cells (left) but does not inhibit the vitamin A uptake activity of COS-1 cells transfected with STRA6 and LRAT (S/L) (right). L, LRAT transfected cells. The activity of S/L cells without MI is defined as 100%. (**D**) MI effectively inhibits HRP endocytosis by WiDr cells (left), but does not inhibit vitamin A uptake activity of WiDr cells (right). The activity of WiDr cells treated with 40 mM retinoic acid but without MI is defined as 100%. (**E**) Vitamin A uptake activity of membrane fractions of human embryonic kidney 293 cells. U, S, L, and S/L denote membranes from untransfected cells, cells transfected with STRA6, LRAT and both, respectively. R, 100× holo-RBP. The activity of S/L cells is defined as 100%. (**F**) RBP remains outside the cell after STRA6-mediated vitamin A uptake from holo-RBP for 8 hours. Absorption spectra of RBP that remains in the supernatant of untransfected cells (left) and cells transfected with STRA6 and LRAT (right) are shown.

culture medium after cellular vitamin A uptake. We purified the RBP in the cell culture medium after cellular uptake of vitamin A and measured the absorption spectrum of the purified RBP (Fig. 4F). A preferential decrease in the vitamin A peak for RBP taken from the medium of STRA6- and LRAT-transfected cells again argued against an endocytosis-based mechanism.

Localization of STRA6 is consistent with its function as the RBP receptor. The tissue localization of STRA6 has been studied previously (*24*). STRA6 is expressed during embryonic development and in the adult brain, spleen, kidney, female genital tract, and testis (and at lower quantities in heart and lung). Furthermore, STRA6 is highly enriched in the RPE in the adult eye (*24*). To further study the localization of STRA6 in ocular organs, we produced peptide antibodies against bovine STRA6 and performed immunohistochemistry on several adult organs. In the RPE, STRA6 is localized to the basolateral membrane (Fig. 5, A to C), a location consistent with its role in interacting with RBP in the choriocapillaris blood (*14*). In contrast to its presence in the endothelial cells of the choriocapillaris (Fig. 5, A and B), STRA6 was expressed in retinal blood vessels, another location of the blood-retina barrier (Fig. 5, D and E). Holo-RBP from retinal blood vessels is a potential source of vitamin A for Müller cells in the retina (*37*). In addition to blood-brain barriers as observed in previous studies (*27*), we found STRA6 expression in astrocyte perivascular endfeet that surround blood vessels negative for STRA6 (Fig. 5F). In accord with National Center for Biotechnology Information (NCBI)'s expressed sequence tag (EST) tissue expression profile for STRA6, we observed STRA6 expression in the placenta (Fig. 5G) and the spleen (Fig. 5H).

Discussion. The existence of an RBP receptor is supported by a large body of evidence. The absence of vitamin A from abundant erythrocytes and serum albumin, which can bind vitamin A, argues against simple diffusion of retinol from the RBP to the cell membrane as the mechanism of vitamin A delivery by RBP (*2*). The fact that free vitamin A can diffuse through membranes is not a good argument against the existence of an RBP receptor. First, nearly all vitamin A in blood is bound to the soluble RBP-TTR complex, which cannot freely pass through cell membranes. Second, there are known examples of molecules that can diffuse through membranes but still depend on membrane transporters to facilitate their transport. For example, prostaglandin can pass the membrane by simple diffusion, but prostaglandin transporter greatly facilitates its transport across membranes (*41*). In addition, there is strong experimental evidence supporting the existence of an RBP receptor that mediates cellular vitamin A uptake (*13*, *23*). This study has identified STRA6 as the RBP receptor. The RBP-STRA6 system represents a small-molecule delivery mechanism that involves an extracellular carrier protein but does not depend on endocytosis. Given the potent biological effects (including toxicity) of vitamin A and its derivatives (*2*, *4*), a controlled release of vitamin A into cells from holo-RBP through STRA6 has an evolutionary advantage over nonspecific diffusion of vitamin A. This mechanism makes it possible to achieve high efficiency and specificity for vitamin A delivery to organs distant from the liver, such as the eye, brain, placenta, and testis.

STRA6's localization in adult brain is consistent with vitamin A's roles in regulating adult brain activities (*7*), such as synaptic plasticity (*35*) and cortical synchronization during sleep (*36*). STRA6 expression in the placenta and testis is consistent with the role of RBP in delivering vitamin A to these tissues (*7*, *42*). STRA6 expression in the spleen and ovaries is consistent with a role of vitamin A in immune regulation and in prevention against infectious disease (*5*, *43*). According to NCBI's EST analysis, STRA6 is also expressed in fetal skin and lung, both known to depend on vitamin A for proper function (*44*, *45*). Undifferentiated human skin keratinocytes were found to have the highest RBP binding activity of any cell or tissue types tested (*18*). The absence or very low expression of STRA6 in the liver (*24*) makes physiological sense. Liver is the major site of production of vitamin A–loaded RBP to deliver vitamin A to peripheral organs. If STRA6 was highly expressed in the liver and absorbed vitamin A from holo-RBP produced by liver itself, a short circuit would be created.

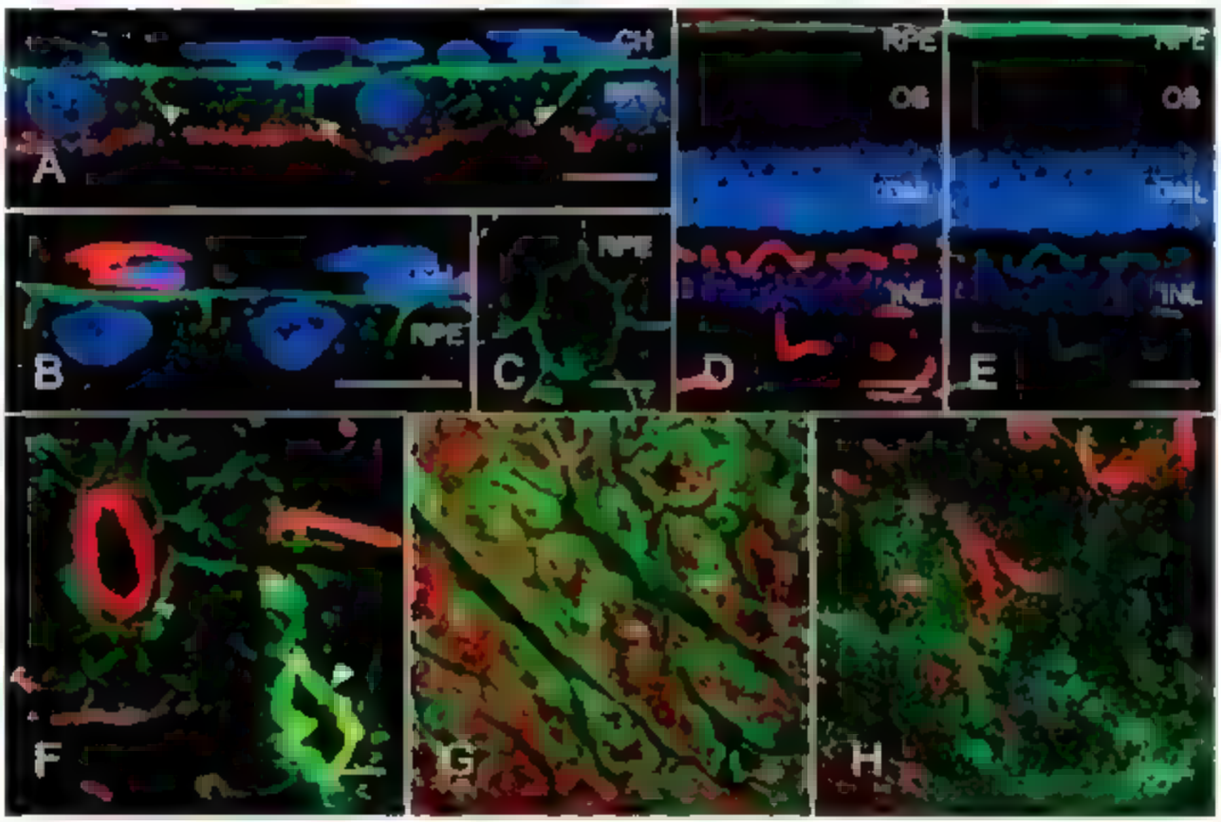


Fig. 5. Immunolocalization of STRA6. Green signal is STRA6. Red signal is peropsin, an apical RPE marker (A) or GSL I–isolectin B4, an endothelial cell marker (B, D, F, G, and H). Blue signal is nuclear stain 4′,6-diamidino-2-phenylindole (DAPI). Scale bars represent 10 μm in (A) to (C) and 40 μm in (D) to (H). CH, choroid; OS, photoreceptor outer segments; ONL, outer nuclear layer; and INL, inner nuclear layer. (**A** and **B**) STRA6 is localized to the basolateral membrane of the RPE. Arrowheads indicate signals on the lateral membrane of RPE cells. (**C**) A tangential section of the RPE layer. (**D** and **E**) STRA6 signals are most enriched in the RPE layer of the eye but are also present in the blood vessels in retina. (E) is an overexposed picture of (D) to show the weaker signals in retina blood vessels. (**F**) Localization of STRA6 in hippocampus. Examples of large blood vessels positive for STRA6 (large arrowhead) and negative for STRA6 (small arrowhead). The STRA6-negative vessel is surrounded by STRA6-positive astrocyte perivascular endfeet. (**G**) Localization of STRA6 in placenta. (**H**) STRA6 is highly expressed in a subset of cells in the spleen.

STRA6 was originally found to be a retinoic acid stimulated gene in cancer cell lines (*24–30*). However, it is possible that there are certain noncancer cell types that respond to retinoic acid to stimulate STRA6 expression. For example, vitamin A combined with retinoic acid increases retinol uptake in the lung in a synergistic manner (*40*), consistent with the ability of retinoic acid to stimulate STRA6 expression. One study found that STRA6 was overexpressed up to 172-fold in 14 out of 14 human colorectal tumors relative to the normal colon tissue (*30*). Thus, the RBP-STRA6 system not only functions as a physiological mechanism of vitamin A uptake but also potentially participates in pathological processes such as insulin resistance (*6*) and cancer.

Note added in proof: A recent human genetic study (*41*) found that mutations in the human *STRA6* gene are associated with widespread birth defects in multiple organ systems. This is consistent with the expression of STRA6 and the diverse functions of vitamin A in embryonic development.

References and Notes

1. J. E. Dowling, *Sci. Am.* **215**, 78 (1966).
2. R. M. Evans, *Harvey Lect.* **90**, 105 (1994).
3. M. Mark, N. B. Ghyselinck, P. Chambon, *Annu. Rev. Pharmacol. Toxicol.* **46**, 451 (2006).
4. D. S. Goodman, in *The Retinoids*, M. B. Sporn, A. B. Roberts, D. S. Goodman, Eds. (Academic Press, New York, 1984), vol. 2, pp. 41–88.
5. R. Blomhoff, M. H. Green, T. Berg, K. R. Norum, *Science* **250**, 399 (1990).
6. Q. Yang *et al.*, *Nature* **436**, 356 (2005).
7. L. Quadro *et al.*, *EMBO J.* **18**, 4633 (1999).
8. A. Sommer, *Nat. Med.* **3**, 1061 (1997).
9. L. Quadro *et al.*, *J. Infect. Dis.* **182** (suppl. 1), 597 (2000).
10. M. W. Seeliger *et al.*, *Invest. Ophthalmol. Vis. Sci.* **40**, 3 (1999).
11. L. Quadro *et al.*, *Endocrinology* **146**, 4479 (2005).
12. N. B. Ghyselinck *et al.*, *Dev. Dyn.* **235**, 1608 (2006).
13. J. Heller, *J. Biol. Chem.* **250**, 3613 (1975).
14. D. Bok, J. Heller, *Exp. Eye Res.* **22**, 395 (1976).
15. L. Rask, P. A. Peterson, *J. Biol. Chem.* **251**, 6360 (1976).
16. C. C. Chen, J. Heller, *J. Biol. Chem.* **252**, 5216 (1977).
17. A. Sivaprasadarao, J. B. Findlay, *Biochem. J.* **255**, 571 (1988).
18. S. Smeland *et al.*, *Biochem. J.* **305**, 419 (1995).
19. A. Sivaprasadarao, J. B. Findlay, *Biochem. J.* **300**, 437 (1994).
20. M. Sundaram *et al.*, *J. Biol. Chem.* **273**, 3336 (1998).
21. P. N. MacDonald, D. Bok, D. E. Ong, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4265 (1990).
22. J. L. Shingleton, M. K. Skinner, D. E. Ong, *Biochemistry* **28**, 9641 (1989).
23. E. Hagen *et al.*, *J. Nutr. Biochem.* **10**, 345 (1999).
24. P. Bouillet *et al.*, *Mech. Dev.* **63**, 173 (1997).
25. S. Ottonello, S. Petrucco, G. Maraini, *J. Biol. Chem.* **262**, 3975 (1987).
26. A. Ruiz *et al.*, *J. Biol. Chem.* **274**, 3834 (1999).
27. M. L. Batten *et al.*, *J. Biol. Chem.* **279**, 10422 (2004).
28. L. Liu, L. J. Gudas, *J. Biol. Chem.* **280**, 40226 (2005).
29. H. L. Monaco, M. Rizzi, A. Coda, *Science* **268**, 1039 (1995).
30. W. Szeto *et al.*, *Cancer Res.* **61**, 4197 (2001).
31. N. L. Mata, R. A. Radu, R. C. Clemmons, G. H. Travis, *Neuron* **36**, 69 (2002).
32. A. Sivaprasadarao, J. B. Findlay, in *Vitamin A in Health and Disease*, R. Blomhoff, Ed. (Marcel Dekker, New York, 1994), pp. 87–117.
33. N. Kanai *et al.*, *Science* **268**, 866 (1995).
34. U. C. Drager, *Sci. STKE* **2006**, pe10 (2006).
35. D. L. Misner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11714 (2001).
36. S. Maret *et al.*, *Science* **310**, 111 (2005).
37. C. B. Stephensen, *Annu. Rev. Nutr.* **21**, 167 (2001).
38. A. Vahlquist, in *Vitamin A in Health and Disease*, R. Blomhoff, Ed. (Marcel Dekker, New York, 1994), pp. 365–424.
39. H. K. Biesalski, D. Nohr, *Mol. Asp. Med.* **24**, 431 (2003).
40. A. C. Ross *et al.*, *J. Lipid Res.* **47**, 1844 (2006).
41. F. Pasutto *et al.*, *Am. J. Hum. Genet.*, in press.

42. We thank E. Wright, G. Travis, K. Philipson, J. Nathans, B. Khakh, T. Vondriska, and B. Ribalet for helpful discussions and K. Philipson, J. Nathans, A. Rattner, G. Travis, and E. Wright for critical reading of the manuscript. Supported by grants from the Stein-Oppenheimer Foundation and the Ruth and Milton Steinbach Foundation (H.S.). R.K. was supported by NIH training grant 5T32EY07026.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1136244/DC1
Materials and Methods
Figs. S1 and S2
References

12 October 2006; accepted 5 December 2006
Published online 25 January 2007;
10.1126/science.1136244
Include this information when citing this paper.

REPORTS

A Common Explosion Mechanism for Type Ia Supernovae

Paolo A. Mazzali,[1,2,3,4]* Friedrich K. Röpke,[1,5] Stefano Benetti,[4] Wolfgang Hillebrandt[1]

Type Ia supernovae, the thermonuclear explosions of white dwarf stars composed of carbon and oxygen, were instrumental as distance indicators in establishing the acceleration of the universe's expansion. However, the physics of the explosion are debated. Here we report a systematic spectral analysis of a large sample of well-observed type Ia supernovae. Mapping the velocity distribution of the main products of nuclear burning, we constrain theoretical scenarios. We find that all supernovae have low-velocity cores of stable iron-group elements. Outside this core, nickel-56 dominates the supernova ejecta. The outer extent of the iron-group material depends on the amount of nickel-56 and coincides with the inner extent of silicon, the principal product of incomplete burning. The outer extent of the bulk of silicon is similar in all supernovae, having an expansion velocity of ≈11,000 kilometers per second and corresponding to a mass of slightly over one solar mass. This indicates that all the supernovae considered here burned similar masses and suggests that their progenitors had the same mass. Synthetic light-curve parameters and three-dimensional explosion simulations support this interpretation. A single explosion scenario, possibly a delayed detonation, may thus explain most type Ia supernovae.

When a white dwarf (WD) composed of carbon and oxygen, accreting mass from a companion star in a binary system, approaches the Chandrasekhar mass [M_{Ch} ≈ 1.38 solar masses ($M_{\odot}$)], high temperature causes the ignition of explosive nuclear burning reactions that process stellar material and produce energy. The star explodes, leaving no remnant, producing a type Ia supernova (SNIa) (*1*). At high stellar material densities, burning yields nuclear statistical equilibrium (NSE) isotopes, in particular radioactive ^{56}Ni, which decays to ^{56}Co and ^{56}Fe, making the SN bright (*2*). At lower densities, intermediate-mass elements (IMEs) are synthesized. Both groups of elements are observed in the optical spectra of type Ia supernovae (SNeIa) (*3*). An empirical relation between an observed quantity, the blue (*B*)-magnitude decline over the first 15 days after maximum luminosity [$\Delta m_{15}(B)$], and a physical quantity, the SN maximum luminosity (*4*), can be used to determine the distance to a SNIa. This method was applied to very distant SNeIa, leading to the discovery of the accelerating universe (*5*, *6*). How the explosion actually proceeds is, however, debated (*7–9*), as is the nature of the progenitor system: Accretion may occur either from a more massive companion (such as a giant) or via the merging of two carbon-oxygen WDs (*10*). This casts a shadow on the reliability of SNeIa as distance indicators, because intrinsically very different explosions may result in the observed correlation.

We derived the distribution of the principal elements in 23 nearby SNeIa (distances ≲ 40 Mpc) with good spectral coverage extending from before maximum to the late nebular phase about 1 year later. The sample (table S1) covers a wide range of light-curve decline rates and includes peculiar objects such as SN2000cx, which violates the luminosity decline-rate relation.

Because of the hydrodynamic properties of the explosion, the expansion velocity of the ejecta is proportional to the radius and serves as a radial coordinate. As the SN expands, deeper layers are exposed. The outer layers, visible in the first few weeks after the explosion, are dominated by IMEs. Because silicon is the most abundant IME, we measured its characteristic

expansion velocity from the blueshift of the absorption core of the strong SiII 6355 Å line in all spectra where it was visible. This velocity decreases with time. Fitting the postmaximum velocity evolution and extrapolating it to the earliest times, when the outermost parts of the ejecta are visible, we derived the outer extension of the bulk of Si. This represents a lower limit of the outer extent of burning. The Si velocity v(Si) is similar in all SNeIa, regardless of their luminosity: v(Si) = 11,900 ± 1300 km s^{-1}.

A few SNe, defined as high-velocity-gradient (HVG) SNe (*11*), are responsible for most of the dispersion. They have a rapidly decreasing v(Si) before maximum and very-high-velocity CaII lines, possibly the result of high-velocity blobs that carry little mass and kinetic energy but cause high-velocity absorption features (HVFs) (*12*) that can abnormally broaden the SiII line profile (*12*, *13*). Although excluding them from the sample decreases the dispersion in v(Si) significantly [v(Si) = 11,300 ± 650 km s^{-1}], we include all of these SNe in our discussion.

The maximum Si velocity thus measured is a conservative estimate. The deep absorption core is produced in layers of high Si abundance. Si is present at higher velocities, indicated by the wavelength of bluest absorption in the earliest spectrum of each SN. However, measuring the bluest absorption velocity yields a large scatter because the earliest spectra have different epochs and may be affected by HVFs (*12*). Our method reliably determines the outer location of the bulk of ME's.

The inner extent of Si, determined from the asymptotic velocity of the SiII 6355 Å line in postmaximum spectra [fig. 1 in (*11*)], is a steep function of Δm_{15} (*B*). The brightest (slowest-declining) SNe have the thinnest Si zones.

The inner ejecta, dominated by NSE elements, are best observed ~1 year after the explosion, when dilation caused by expansion makes the SN behave like a nebula, exposing the deepest layers. Collisions with the fast particles produced by the decay of ^{56}Ni→ ^{56}Co→^{56}Fe heat the gas, which cools, emitting radiation mostly in forbidden lines.

We modeled the nebular spectra using a code that computes line emission, balancing heating and cooling in nonlocal thermodynamic equilibrium (NLTE) (*14*), including density and abundance stratification. We adopted the density-velocity distribution of the standard, one-dimensional, M_{Ch} explosion model W7 (*1*) (fig. S1). In the nebular phase, the gas is transparent, and line emissivity depends on the mass of the emitting ion. Accurately estimating this mass requires determining the ionization state of the gas. Forbidden lines of FeII and FeIII dominate SNIa nebular spectra, reflecting the high abundance of NSE material. Fe is mostly the product of ^{56}Ni decay, which provides heating. The stable, neutron-rich isotopes ^{54}Fe and ^{58}Ni do not contribute to heating but do contribute to cooling, because they also emit forbidden lines. Their presence affects the ionization balance. Both are mostly produced deep in the WD, at the highest densities. The ^{54}Fe nebular lines have wavelengths indistinguishable from those of ^{56}Fe.

We determined the mass and distribution in velocity of the Fe isotopes (and thus of ^{56}Ni), simultaneously fitting the ratio of the two strongest Fe emissions (fig. S1). One, near 4700 Å, includes both FeII and FeIII lines, whereas the other, near 5200 Å, is only due to FeII. A low upper limit to the mass of ^{58}Ni is set by the absence of strong emission lines (in particular at 7380 Å).

In the deepest layers (Fig. 1), all SNe contain ~0.1 to 0.3 $M_\odot$ of stable NSE isotopes, with a large scatter and no dependence on Δm_{15} (*B*) [see also (*15*)].

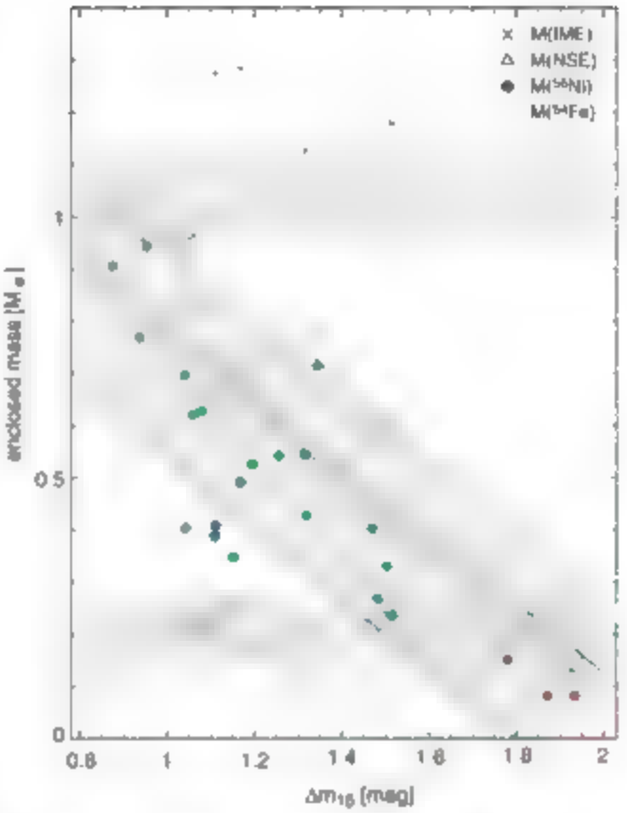


Fig. 1. The Zorro diagram. Distribution of the principal isotopic groups in SNeIa. The enclosed mass (linked to velocity via the W7 explosion model) of different burning products is shown versus decline-rate parameter Δm_{15} (*B*) (a proxy for SN luminosity). Individual SNe are colored according to their velocity evolution (*11*): HVG, blue; low velocity gradient (LVG), green; and faint, red. Open circles indicate the mass of stable ^{54}Fe + ^{58}Ni for each SN, solid circles indicate that of ^{56}Ni, and open triangles indicate the sum of these (total NSE mass). Crosses show the sum of NSE and IME mass, indicating the total mass burned. The IME mass is the difference between crosses and triangles. ^{54}Fe and ^{58}Ni are found in roughly constant amounts in the deepest parts of all SNe, irrespective of luminosity: M(stable NSE) = 0.20 ± 0.05 $M_\odot$ (lower horizontal shaded area). The ^{56}Ni mass determines the SN luminosity. It correlates with Δm_{15} (*B*): $M(^{56}\mathrm{Ni}) = 1.34 - 0.67\,\Delta m_{15}$ (*B*) [$M_\odot$] with rms dispersion 0.13 $M_\odot$ (lower diagonal shaded area). The total NSE mass correlates with Δm_{15} (*B*) better than $M(^{56}\mathrm{Ni})$: M(NSE) = 1.55 − 0.69Δm_{15} (*B*) [$M_\odot$], rms dispersion 0.09 $M_\odot$ (upper diagonal shaded area). IMEs lie mostly outside the iron-group zone. The outer Si velocity is similar for all SNe except HVG SNe. The mass enclosed by IMEs represents the total burned mass. When all SNe are included, the average value is 1.05 ± 0.09 $M_\odot$ (upper horizontal shaded area). Excluding HVG SNe, the value is 1.01 ± 0.05 $M_\odot$ (horizontal line). Both values are independent of Δm_{15} (*B*). For a version of this plot with SN names and a bar diagram, see figs. S2 and S3.

As expected, the ^{56}Ni mass correlates inversely with Δm_{15} (*B*), ranging from 0.9 $M_\odot$ for the slowest-declining (most luminous) SNe to 0.1 $M_\odot$ for the fastest-declining (dimmest) ones. The root mean square (rms) dispersion is 0.13 $M_\odot$, but SNe with intermediate decline rates [Δm_{15} (*B*) ~ 1.05 to 1.5 magnitudes] show variations of almost a factor of 2 for the same value of Δm_{15} (*B*). These SNe could cause scatter about the mean luminosity decline-rate relation (*16*).

Once the contributions of ^{56}Ni, ^{54}Fe, and ^{58}Ni are added together to evaluate the total NSE mass, the dispersion decreases to 0.09 $M_\odot$. If SNe with different amounts of ^{56}Ni, and thus presumably different temperatures, but similar NSE content have similar Δm_{15} (*B*), it is likely that abundances (*17*) rather than temperature (*18*) primarily determine the opacity and light-curve shape.

The outer velocity of the NSE region, determined from the width of the Fe lines, correlates with SN luminosity (*19*). It coincides with the innermost Si velocity, marking the transition from complete to incomplete burning. The difference between these two velocities, Δv(·) = 650 ± 900 km s^{-1}, is consistent with zero. This results from two different methods applied to data obtained almost 1 year apart.

Thus, whereas the mass of ^{56}Ni, and consequently the SN luminosity, can differ significantly

[1]Max-Planck Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, 85741 Garching, Germany. [2]Department of Astronomy, School of Science, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan. [3]Research Center for the Early Universe, School of Science, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan. [4]Istituto Nazionale di Astrofisica–Osservatorio Astronomico di Trieste, Via Tiepolo 11, I-34131 Trieste, Italy. [5]Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA. [6]Istituto Nazionale di Astrofisica-Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio, 2, I-35122 Padova, Italy.

*To whom correspondence should be addressed. E-mail: mazzali@mpa-garching.mpg.de

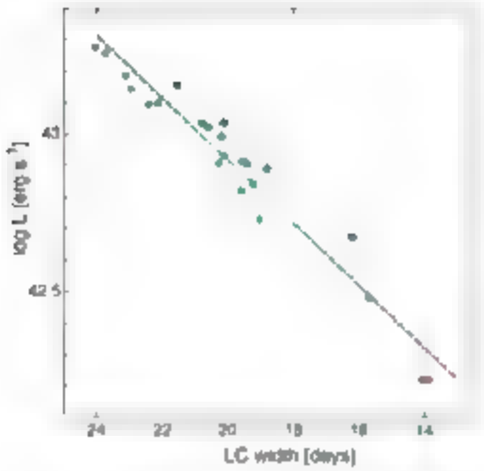


Fig. 2. Observed and synthetic luminosity–decline-rate relations for the SNe in our sample. Colors indicate velocity evolution as in Fig. 1. The peak luminosity L was computed from $M(^{56}\mathrm{Ni})$ as $L = 2 \times 10^{43}\ M(^{56}\mathrm{Ni})$ (*28*). Observed bolometric light-curve (LC) widths were obtained from observed rise and decline times (*29*) as $T = (t_{-1/2} + t_{+1/2})$. When $t_{-1/2}$ is missing, it was estimated from the relation between $L_{1/2}$ and Δm_{15} (*8*) derived from the other SNe. The dashed line is a linear regression between luminosity and observed LC width for the seven SNe common to our sample and that of (*29*). Synthetic bolometric LC widths were obtained assuming that: (i) The LC width τ depends on ejected mass M_{ej}, kinetic energy E_k, and opacity κ as $t \propto \kappa^{\frac{1}{2}} E_k^{-\frac{1}{4}} M_{ej}^{\frac{3}{4}}$ (*21*). (ii) For all SNe, $M_{ej} = M_{Ch}$. The W7 (*1*) density-velocity distribution was used. (iii) The explosion kinetic energy depends on the burning products: $E_k = [1.56\ M(^{56}\mathrm{Ni}) + 1.74\ M(\mathrm{stable\ NSE}) + 1.24\ M(\mathrm{IME}) - 0.46] \times 10^{51}$ erg (*17*). (iv) Opacity is mostly due to line absorption (*30*). Accordingly, NSE elements contribute much more than IMEs, because their atomic level structure is more complex. The opacity was therefore parametrized according to the abundances of different species, $\kappa \propto M(\mathrm{NSE}) + 0.1\ M(\mathrm{IME})$ (*16*). To compare the parametrized LC widths τ to observed values T, a scale factor $x = T/\tau$ was computed for each of the seven SNe common to our sample and that of (*29*). The average factor, $x = 24.447$, was used to scale all SNe. Dots show the individual SNe. The continuous line is the linear regression between luminosity and synthetic LC width for our sample of 23 SNeIa. For a version of this plot with SN names, see fig. S4.

other characteristics of SNeIa are remarkably homogeneous. In particular, the narrow dispersion of the outer Si velocity indicates a similar extent of thermonuclear burning in all SNeIa. SNe that produce less ^{56}Ni synthesize more IMEs.

The simplest interpretation of these seemingly antithetical results is that thermonuclear burning consumes similar masses in all SNeIa. We explore whether this scenario is consistent with the M_{Ch} model (*10*). Applying the density-velocity structure of model W7 (*1*), we transform v(Si) to mass (Fig. 1). We find that the outer shell of Si encloses a mass of at least $1.05 \pm 0.09\ M_\odot$ ($1.01 \pm 0.05\ M_\odot$, excluding HVG SNe), independently of Δm_{15} (*8*). This is a lower limit to the burned mass.

The light-curve width τ is related to Δm_{15} (*8*) (*20*) and depends on the ejected mass M_{ej}, the kinetic energy E_k, and the opacity κ as $\tau \propto \kappa^{\frac{1}{2}} E_k^{-\frac{1}{4}} M_{ej}^{\frac{3}{4}}$ (*21*). We tested our assumption by computing parametrized light-curve widths and comparing them to observed values. The resulting luminosity–decline-rate relation (Fig. 2) is very tight and practically identical to the observed one. This result supports our opacity parametrization, corroborating our hypothesis that the mass burned is similar in all SNeIa. Given its weak dependence on E_k, the light-curve width is not much affected if more of the outer part of the WD is burned to IMEs, as these do not contribute much to the opacity.

If SNeIa burn a similar mass, the progenitor mass is also likely to be the same, namely M_{Ch}. Because the outcome of the burning depends essentially on fuel density, a variation in iron-group elements production in M_{Ch} WDs requires different WD expansion histories. This in turn depends on the details of the burning. Once the WD reaches M_{Ch}, a thermonuclear flame is ignited near the center. The flame must start as a subsonic deflagration (*21*), mediated by microphysical transport and accelerated by turbulence. As it propagates outwards, it could undergo a deflagration-to-detonation transition (DDT) and continue as a shock-driven supersonic detonation wave, in a so-called delayed detonation (*22*). This constitutes the most extreme explosion scenario admissible, exploring the limits of M_{Ch} explosions.

We modeled the explosion using a three-dimensional level-set approach (*23*). The ignition of the deflagration flame was treated as a stochastic process generating a number of ignition spots placed randomly and isotropically within 130 km of the WD center (*24*).

What would cause the DDT is unclear. We assumed that it occurs as turbulence penetrates the internal flame structure, the onset of the so-called distributed burning regime (*25*). This happens at low fuel densities, after some WD pre-expansion in the deflagration phase. The detonation is triggered artificially where the chosen DDT criterion is first satisfied, typically near the outer edges of the deflagration structure (Table 1 shows DDT parameters).

Three simulations, with 800, 20, and 5 ignition spots, termed D800, D20, and D5, respectively, were performed on a moving cellular Cartesian grid (*26*) comprising the full star. Model D800, with its dense distribution of ignition points, exhausts the carbon-oxygen fuel at the WD center almost completely in the deflagration phase. The energy release quickly expands the star, and the subsequent detonation mainly transforms low-density outer material to IMEs. In contrast, the few ignition spots of model D5 consume little material during the deflagration, leaving more fuel at high densities, which is converted mostly to NSE isotopes in the vigorous detonation phase. Model D20 provides an intermediate case.

Table 1. Results and DDT parameters of parametrized three-dimensional delayed detonation models.

Model	E_k^{asympt} (10^{51} erg)	M(NSE) ($M_\odot$)	M(IME) ($M_\odot$)	t_{DDT} (s)	ρ_{DDT} (10^7 g cm^{-3})
D800	1.004	0.638	0.547	0.675	2.40
D20	1.237	0.833	0.435	0.724	1.92
D5	1.524	1.141	0.220	0.731	1.33

The ejecta compositions of the model explosions agree grossly with the results derived from the spectra. The NSE mass produced ranges from 0.638 to 1.141 $M_\odot$. The weakest explosions result from optimal burning in the deflagration phase (*26*), and subluminous SNeIa are not reached in our parametrization. However, the conditions for DDT need further investigation. Eventually these events may be explained within a single framework. Some extremely luminous SNeIa may come from very rapidly rotating WDs whose mass exceeds M_{Ch} (*27*), but these are rare.

The distribution derived from observations of burning products inside SNeIa could result from the variation of a single initial parameter, the flame ignition configuration, in M_{Ch} delayed detonations. The luminosity–decline-rate relation can be reproduced using this distribution and a simple opacity parametrization. Our results support the M_{Ch} scenario for most SNeIa, adding confidence to their use as distance indicators.

References and Notes

1. K. Nomoto, F.-K. Thielemann, K. Yokoi, *Astrophys. J.* **286**, 644 (1984).
2. M. J. Kuchner, R. P. Kirshner, P. A. Pinto, B. Leibundgut, *Astrophys. J.* **426**, L89 (1994).
3. D. Branch, J. B. Doggett, K. Nomoto, F.-K. Thielemann, *Astrophys. J.* **294**, 619 (1985).
4. M. M. Phillips, *Astrophys. J.* **413**, L105 (1993).
5. A. G. Riess *et al.*, *Astron. J.* **116**, 1009 (1998).
6. S. Perlmutter *et al.*, *Astrophys. J.* **517**, 565 (1999).
7. M. Reinecke, W. Hillebrandt, J. C. Niemeyer, *Astron. Astrophys.* **391**, 1167 (2002).
8. V. N. Gamezo, A. M. Khokhlov, E. S. Oran, *Phys. Rev. Lett.* **92**, 211102 (2004).
9. T. Plewa, A. C. Calder, D. Q. Lamb, *Astrophys. J.* **612**, L37 (2004).
10. W. Hillebrandt, J. C. Niemeyer, *Annu. Rev. Astron. Astrophys.* **38**, 191 (2000).
11. S. Benetti *et al.*, *Astrophys. J.* **623**, 1011 (2005).
12. P. A. Mazzali *et al.*, *Astrophys. J.* **623**, L37 (2005).
13. M. Tanaka, P. A. Mazzali, K. Maeda, K. Nomoto, *Astrophys. J.* **645**, 470 (2006).
14. P. A. Mazzali, K. Nomoto, F. Patat, K. Maeda, *Astrophys. J.* **559**, 1047 (2001).

15. S. E. Woosley, D. Kasen, S. Blinnikov, E. Sorokina, preprint available at http://arxiv.org/ps/astro-ph/0609562 (2006).
16. P. A. Mazzali, P. Podsiadlowski, *Mon. Not. R. Astron. Soc.* **369**, 19 (2006).
17. P. A. Mazzali *et al.*, *Ap. J.* **547**, 988 (2001).
18. D. Kasen, S. E. Woosley, preprint available at http://arxiv.org/ps/astro-ph/0609540 (2006).
19. P. A. Mazzali, E. Cappellaro, I. J. Danziger, M. Turatto, S. Benetti, *Astrophys. J.* **499**, 149 (1998).
20. G. Goldhaber *et al.*, *Astrophys. J.* **558**, 359 (2001).
21. W. D. Arnett, *Astrophys. J.* **253**, 785 (1982).
22. A. M. Khokhlov, *Astron. Astrophys.* **245**, 114 (1991).
23. I. Golombek, J. C. Niemeyer, *Astron. Astrophys.* **438**, 611 (2005).
24. S. E. Woosley, S. Wunsch, M. Kuhlen, *Astrophys. J.* **607**, 921 (2004).
25. J. C. Niemeyer, S. E. Woosley, *Astrophys. J.* **475**, 740 (1997).
26. F. K. Roepke, W. Hillebrandt, J. C. Niemeyer, S. E. Woosley, *Astron. Astrophys.* **448**, 1 (2006).
27. D. A. Howell *et al.*, *Nature* **443**, 308 (2006).
28. M. Stritzinger, S. Leibundgut, *Astron. Astrophys.* **431**, 423 (2005).
29. G. Contardo, B. Leibundgut, W. D. Vacca, *Astron. Astrophys.* **359**, 876 (2000).
30. A. W. A. Pauldrach *et al.*, *Astron. Astrophys.* **312**, 525 (1996).

31. We thank E. Pian and D. Sauer for help with data analysis. This work was partly supported by the European Union's Human Potential Programme under contract HPRN-CT-2002-00303, "The Physics of Type Ia Supernovae."

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5813/825/DC1
Figs. S1 to S5
Table S1
References

12 October 2006; accepted 12 December 2006
10.1126/science.1136259

Coding/Decoding and Reversibility of Droplet Trains in Microfluidic Networks

Michael J. Fuerstman,[1] Piotr Garstecki,[2]* George M. Whitesides[1]*

Droplets of one liquid suspended in a second, immiscible liquid move through a microfluidic device in which a channel splits into two branches that reconnect downstream. The droplets choose a path based on the number of droplets that occupy each branch. The interaction among droplets in the channels results in complex sequences of path selection. The linearity of the flow through the microchannels, however, ensures that the behavior of the system can be reversed. This reversibility makes it possible to encrypt and decrypt signals coded in the intervals between droplets. The encoding/decoding device is a functional microfluidic system that requires droplets to navigate a network in a precise manner without the use of valves, switches, or other means of external control.

The use of microfluidic devices provides a means to study both nonlinear (periodic and chaotic) and reversible behaviors of flows of fluids, although no microfluidic system has simultaneously demonstrated both of these classes of dynamics. At low Reynolds number (Re), the behavior of fluids typically can be reversed—as if the system were going backward in time—by reversing the direction of the forces applied to the system. Recent work has detailed microfluidic systems, which demonstrate nonlinear behaviors such as periodic or chaotic dynamics, that are either inherently irreversible or have not been demonstrated to be reversible (*1*–*3*). Here, we describe a microfluidic device that displays reversible nonlinear dynamics. In this system, droplets of water-based ink moving in hexadecane arrive at a T intersection, where they select one of two paths that form a loop by reconnecting downstream. The binary choices that the droplets make cause the resistance to flow through each branch to evolve in time in a nonlinear fashion: that is, the droplets in each branch influence the choice of subsequent droplets between branches by modifying the rates of flow of liquid through the channels. This behavior results in nonlinear dynamics: droplets select paths with controllable periodicity, and in some instances they choose paths in an aperiodic fashion. Whereas the characteristically linear flow of both liquids through the microchannels allows the dynamics to be reversible. We believe that this system can serve as a platform for studying the transition in nonlinear systems from reversible to irreversible dynamics. The reversible, nonlinear dynamics that the device demonstrates are also useful for "lab-on-a-chip" applications—systems that perform physical, chemical, or biological functions on millimeter- to centimeter-scale platforms. For example, we used the device to process a signal represented by sequences of droplets—by encoding and then decoding it—using only the pressure-driven flow of liquids.

Flows at the small length scales (10 to 100 μm) that are characteristic of microfluidic systems typically occur at low Re and thus are dominated by viscosity; inertia plays only a marginal role (*4*). Viscosity-dominated flows are governed by equations of motion that are linear in the velocity of the fluid. Classic experiments by Taylor (*5*) demonstrated how low-Re flows can be reversed. A drop of dye added to a viscous fluid between two cylinders was stretched by rotating the inner one and subsequently reconstituted by reversing the direction of rotation. Recently, Pine *et al.* have investigated a similar system in which colloidal particles in the bulk fluid took the place of the dye in Taylor's experiments (*6*). They demonstrated that the reversibility of the dynamics of the system depended critically on the number of nonlinear events—that is, collisions between particles—that took place during rotation. In both of these systems and in ours, the motion of the carrier fluid can be approximately described by the reversible Stokes equation (*7*).

Microfluidic systems have recently been used to generate multiphase flows where the fluids include suspended droplets (*8*–*10*) and bubbles (*11*–*13*) with substantial control over their size, volume fraction in the carrier fluid, and frequency of production (*14*, *15*). The interfacial stresses present in these multiphase systems introduce nonlinearities into the equations of flow, even at low values of Re. These nonlinearities are weak, however, compared with the linear contribution to the dynamics (*16*), and therefore do not affect the reversibility of the movement of bubbles or droplets through a microchannel.

We intentionally introduce an additional strongly nonlinear component into the interactions that govern the motion of droplets. Fig. 1A shows the device we use. The nonlinearity arises from the binary choice that the droplets make at the T intersection. Each droplet that enters this intersection chooses the branch of the loop through which the hexadecane flows more rapidly—equivalently, the path characterized by the smaller fluidic resistance (*17*). Because the system operates at a capillary number ($< 10^{-1}$) that is small enough that the droplets remain intact as they move through the T intersection (*18*), the process of choosing a path amplifies the differences in the rates of flow of hexadecane through the branches of the loop into a binary value: a droplet of aqueous solution is either present in or absent from a branch.

When a droplet moves through a branch of the loop, it increases the resistance to flow in that channel. The droplet consequently decreases the rate of flow of hexadecane through the branch it occupies and increases (at a constant rate of flow) the rate of flow through the other branch (*19*–*21*). Because the next droplet to enter the T intersection also enters the channel through which the hexadecane flows more rapidly, the choice that one droplet makes influences the choice of the next droplet arriving at the junction. This feed-

[1]Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA. [2]Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, 01 224 Warsaw, Poland.

*To whom correspondence should be addressed. E-mail: gwhitesides@gmwgroup.harvard.edu (G.M.W.); garst@ichf.edu.pl (P.G.)

back between successive droplets results in a nonlinear system that exhibits bifurcations and irregular aperiodic behavior. The nonlinearity in the operation of the system derives from two sources: (i) the time evolution of the fluidic resistances of the two arms of the loop and (ii) the nonlinear transformation of the intervals between droplets entering the loop and those between droplets leaving it. Remarkably, as a result of the precise confinement of the source of the nonlinearity to isolated events, this strongly nonlinear device demonstrates reversibility, a characteristic typical of, but not limited to, linear systems.

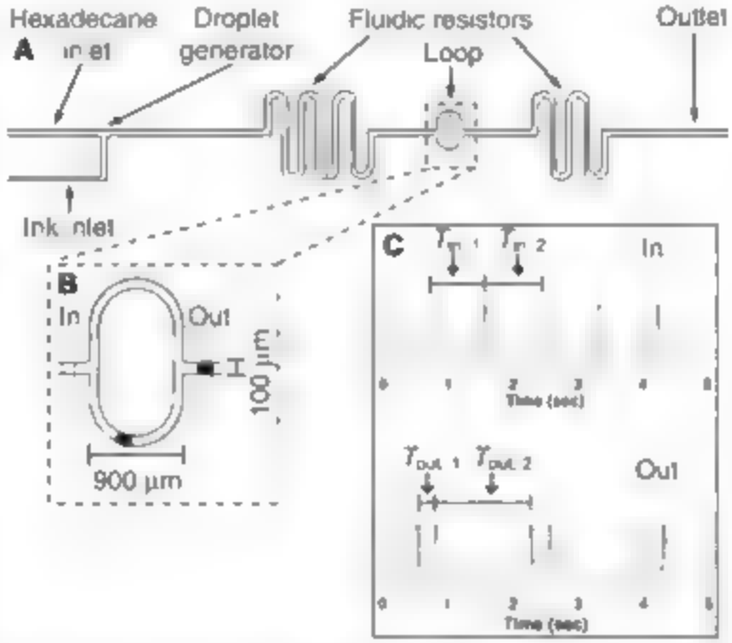


Fig. 1. The microfluidic device. (**A**) A generalized schematic diagram of the microchannel network. Hexadecane containing a surfactant (3% Span-80 by mass) squeezed off droplets of aqueous ink or dye in the T junction. The droplets proceeded to the loop, where they took one of the two paths. (**B**) An optical micrograph of the loop showing one droplet in the lower branch and one droplet in the outlet channel. (**C**) Plots showing the intervals between droplets as they moved through the In and Out windows marked in (B). A spike indicates the presence of a droplet. The upper plot shows uniform time intervals ($T_{in,1}$ and $T_{in,2}$) between droplets as they reached the loop. The lower plot depicts the two different time intervals ($T_{out,1}$ and $T_{out,2}$) that separated droplets as they emerged from the loop when the system operated in period-2 mode.

The basic microfluidic system we use comprises a droplet generator, a section in which a channel branches into a loop and then reconnects, and fluidic resistors (long sections of channel either folded into a serpentine geometry for compactness or wound into a spiral). Figure 1A sketches the general form of the network. Pressure applied to a reservoir of hexadecane that contains 3% by mass Span 80, a nonionic surfactant, and a reservoir of an aqueous solution of ink or dye, creates droplets at constant intervals of time at a T junction. The droplets then move through a fluidic resistor, which damps the changes in pressure at the T junction caused by droplets navigating through the loop and ensures that the bubbles are produced at a constant frequency. After the resistor, the droplets reach a loop that has asymmetric branches. In the system shown in Fig. 1B, the longer branch is 1.98 mm long and the shorter 1.78 mm. Finally, the droplets travel through another fluidic resistor and out of the system.

We monitored the times at which droplets enter and exit the loop at the positions marked in Fig. 1B using a high-speed charge-coupled-device camera. The behavior of the system depends on the intervals of time that separate the droplets as they arrive at the entrance to the loop. For sufficiently large intervals between drops (t > 1.04 s), only one droplet passes through the loop at a time. As the intervals between droplets decreases to 0.90 s, a droplet arrives at the junction while the previous one still occupies the shorter branch. The increase in the fluidic resistance of the shorter branch due to the presence of the droplet is sufficient to divert every second droplet into the longer branch. Because the rates of flow through each of the branches differ, the intervals between droplets exiting the loops repeat in sequences of a longer interval followed by a shorter one (Fig. 1C); we refer to this regime as "period-2." For 0.90 > t > 1.04 s, the system exhibits period-3 and period-4 behavior.

As we decrease the intervals between droplets entering the loop, the system displays a series of bifurcations to higher-order periodic and aperiodic behaviors. To characterize the periodicities, we construct Poincaré maps, which are plots of the time interval between the nth and n + 1st droplet versus the time interval between the nth and n − 1st droplet (Fig. 2A). The two plots on the left (a and c) correspond to the intervals between droplets before they reach the loop for two different values of the mean interval between droplets (0.679 and 0.591 s for sets i and ii). The right-hand plot (b) in set i depicts the system demonstrating period-3 behavior; the 33 points on this plot fall into three clusters. The right-hand plot in set ii (d) shows the system behaving in an aperiodic manner: the 42 intervals measured do not cluster on the Poincaré map.

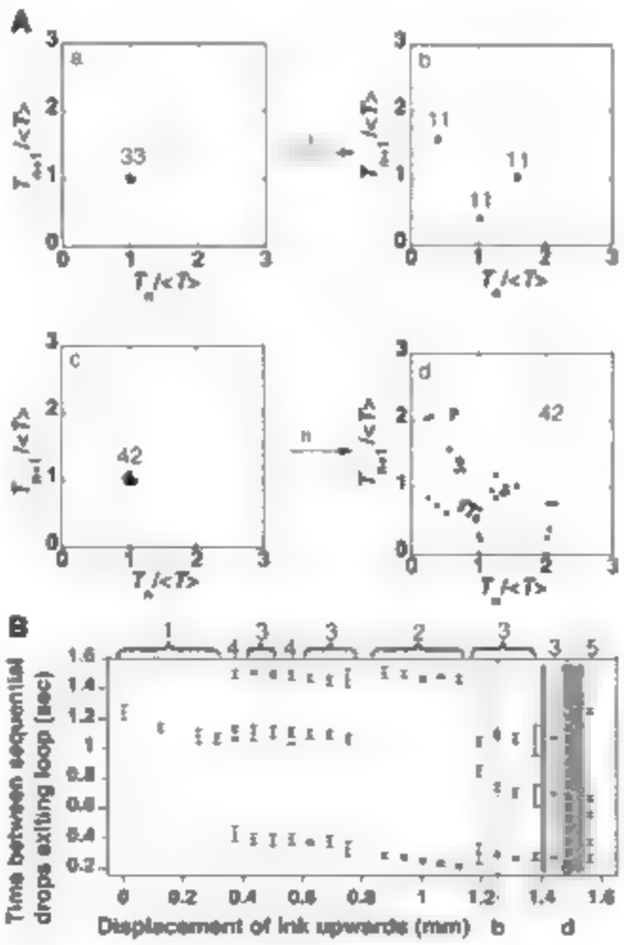


Fig. 2. Periodic and aperiodic behavior. (**A**) Poincaré maps of the system shown in Fig. 1 in (i) period-3 mode and (ii) an aperiodic mode. The maps plot the n + 1st interval (T_{n+1}) versus the nth interval (T_n); the intervals are normalized by the mean of the intervals in each set of data (<T>). The left-hand maps (a and c) show one cluster and indicate that the intervals between drops were uniform as they reached the loop. The number of clusters in map b reveals the trimodal periodicity of the system. Map d shows no discernable clustering and thus denotes aperiodic behavior. The numbers next to the clusters in plots a to c denote the number of data points in each cluster, whereas the number in plot d shows the number of data points in the entire plot. (**B**) A bifurcation diagram for the system. As the displacement (the height of the ink) increased, the periodicity of the system changed. The numbers above the chart denote the periodicity demonstrated for each range of the height of the reservoir of ink. Gray bars denote areas where the system behaved aperiodically. The states of the system as represented by Poincaré maps b and d are marked on the bifurcation diagram.

Figure 2B uses a bifurcation diagram to summarize the periodicities we observe for the one-loop system to the extent that our experimental setup allows us to probe. We fix the pressure of the nitrogen applied to the reservoirs of hexadecane and ink and then raise the vial containing the ink in steps of 250 μm using a micropositioner. Raising this vial increases the pressure applied to the ink due to gravity, increases its rate of flow, and thus decreases the intervals between droplets entering the loop. The resulting periodicities of the paths selected by the droplets change as we vary the pressure applied to the ink. We observe period-1, -2, -3, -4 and -5 behavior, as well as regimes where the droplets select branches aperiodically. The observation of a bifurcation cascade is typical of many nonlinear systems.

To determine whether the dynamics of the system are reversible, we prepare a system containing one loop followed by a spiraling outlet channel ~9.5 cm long (22) in which we store the sequence of droplets that exit the loop. We set the pressure applied to the hexadecane and the aqueous solution of dye so that the device operates in period-7 mode for a sufficiently long period of time that the spiral channel fills with droplets in a period-7 configuration. We then reverse the direction of flow through the system by toggling two three-way valves. The valves switch the applied pressure from the hexadecane supply to the outlet of the device (*23*).

Figure 3A shows a time-space diagram of droplets traveling forward and backward through the system. As the system initially evolves in time (before the reversal of the direction of flow), the droplets move from left to right in the channel (the loop is positioned in the center of the plot, accounting for the blank space). We plot the position of each droplet within the channel versus time as the system evolves. The bands on the left side of the plot are evenly spaced, the uniform spacing indicating that the droplets enter the left side of the loop separated by constant intervals of time (the standard deviation of the intervals is 1.2% of the mean interval). The bands on the right side of the time-space plot show that the droplets exit the right side of the loop in intervals that repeat in groups of seven. The Poincare map on the right (a set 1 of Fig. 3B) confirms that the system operates in a period-7 mode while the continuous liquid flows forward through the system. When we reverse the direction of flow through the channels, the droplets enter the right side of the loop in period-7 mode and emerge from the left side in a stream of uniform periodicity. The standard deviation of the intervals increases, however, to 7.5% of the mean time interval after we reversed the direction of flow. The spiral was large enough to hold ~10 periods of seven droplets each. The system demonstrates reversibility over five period-7 groups of droplets but fails to reverse the period-7 behavior to period-1 behavior for longer sequences. We believe that the failure in the reversibility of the system after five periods and the increase in the standard deviation of the intervals are caused by the droplets moving closer together as they are stored in the inner ring of the spiral.

The reversibility of the dynamics of a system depends on the insensitivity of those dynamics to perturbations. The scatter in the right-hand Poincare map shown in Fig. 3C (ii) demonstrates that the system is subjected to perturbations upon reversal of the direction of flow of hexadecane. Despite these perturbations, the droplets still exit the loop in a period-1 fashion. The dynamics therefore preserve the coherence of the signal.

The device with one loop that we described previously demonstrates reversibility upon reversal of the direction of flow. We found that period-2 sequences, however, are reversible if the droplets are sent forward through a second loop identical to, and downstream of, the first one (Fig. 4, A and B). This property of reversibility without requiring a reversal of the direction of flow is due to backward-forward symmetry. The period-2 sequence ABABAB (where A denotes a longer interval and B a shorter one) looks exactly the same whether read in the forward or reverse direction, with a B always following an A.

We use the two-loop device to encode and decode a signal comprising the intervals of time between droplets. We use the device in Fig. 4A to process an analog signal and the device in Fig. 4B to process a digital one, although both devices are capable of processing either type of signal. To encode the analog signal, we modulate the pressure applied to the reservoir of ink about the

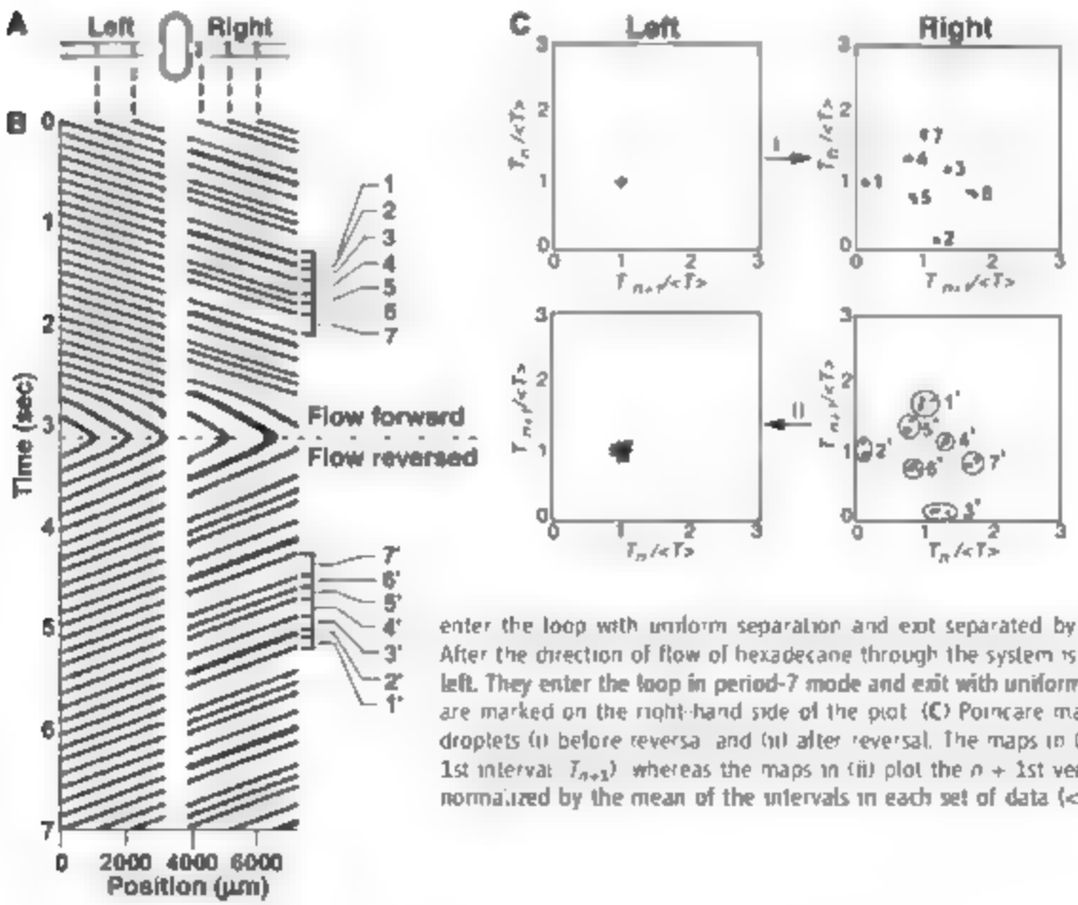


Fig. 3. Reversing the behavior of the system. (**A**) An optical micrograph of droplets moving through the channel on either side of the loop. We stretched the width of the channel relative to its length by a factor of nine, using Adobe Photoshop, to make the droplets more easily visible. The initial positions of the droplets in the optical micrograph correspond to the points at the top of the plot in (**B**) to which dashed lines are drawn. The loop is drawn schematically and is not part of the optical micrograph. (**B**) A plot of the position of droplets in the system as a function of time. Each horizontal slice of this plot represents a snapshot of the system. The black areas signify the presence of a droplet, and the white parts depict the absence of droplets. The droplets initially move from left to right, they enter the loop with uniform separation and exit separated by a repeating sequence of seven intervals. After the direction of flow of hexadecane through the system is reversed, the droplets move from right to left. They enter the loop in period-7 mode and exit with uniform separation. The seven repeating intervals are marked on the right-hand side of the plot. (**C**) Poincare maps showing the periodicity of the train of droplets (i) before reversal and (ii) after reversal. The maps in (i) plot the nth interval (T_n) versus the n + 1st interval (T_{n+1}), whereas the maps in (ii) plot the n + 1st versus the nth. In all maps, the intervals are normalized by the mean of the intervals in each set of data ($\langle T\rangle$).

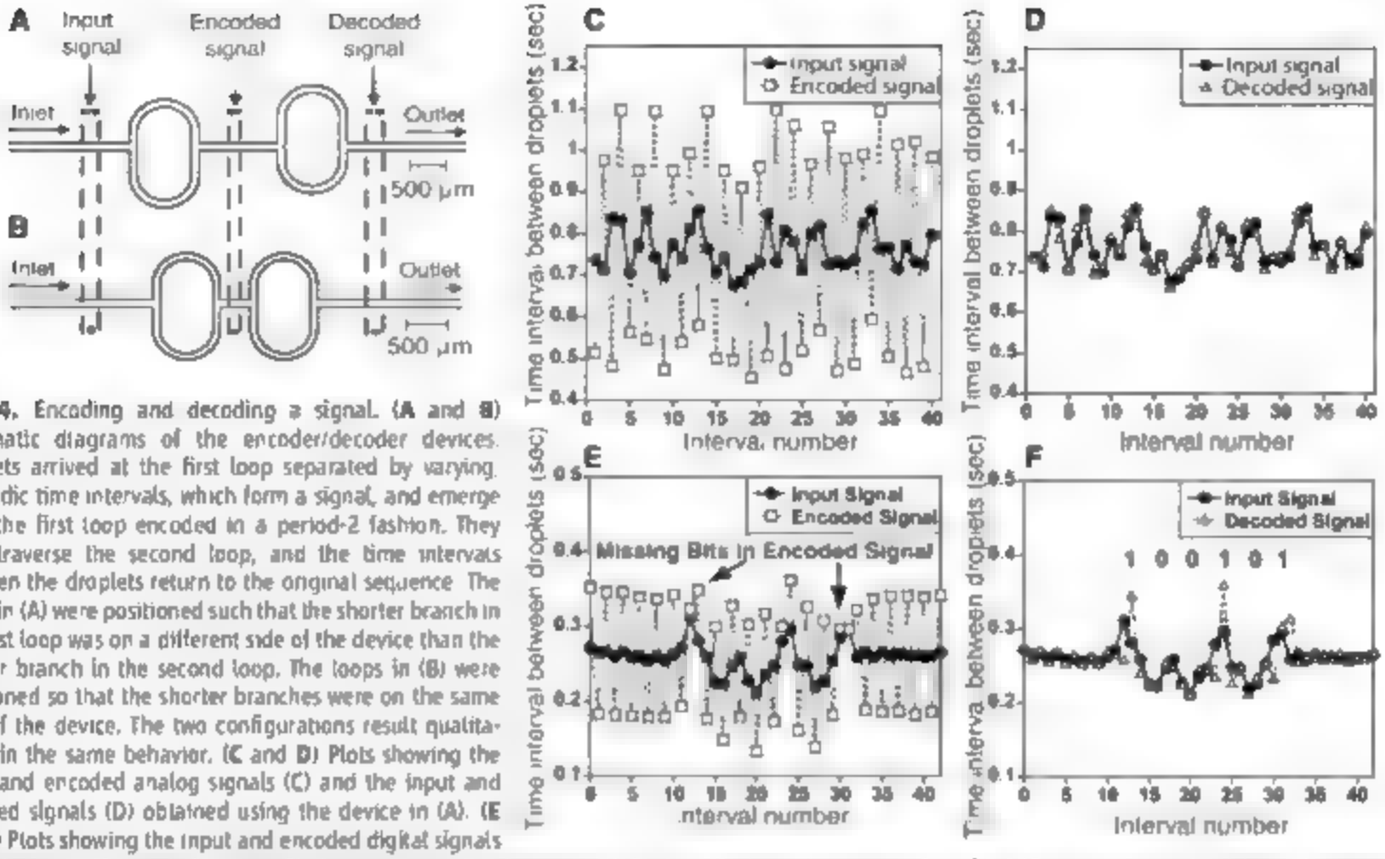


Fig. 4. Encoding and decoding a signal. (**A** and **B**) Schematic diagrams of the encoder/decoder devices. Droplets arrived at the first loop separated by varying, aperiodic time intervals, which form a signal, and emerge from the first loop encoded in a period-2 fashion. They then traverse the second loop, and the time intervals between the droplets return to the original sequence. The loops in (A) were positioned such that the shorter branch in the first loop was on a different side of the device than the shorter branch in the second loop. The loops in (B) were positioned so that the shorter branches were on the same side of the device. The two configurations result qualitatively in the same behavior. (**C** and **D**) Plots showing the input and encoded analog signals (C) and the input and decoded signals (D) obtained using the device in (A). (**E** and **F**) Plots showing the input and encoded digital signals (E) and input and decoded signals (F) obtained using the device in (B). Each peak above the y-axis value of 0.280 s represents a 1, while each valley below 0.225 s represents a 0. The digital signal 100101 is equivalent to the decimal number 37. The arrows in (E) show the bits in the input signal that do not appear in the encoded signal but reappear in the decoded signal in (F).

value corresponding to the midpoint of the region characterized by period-2 behavior by adjusting the height of the reservoir of ink between ~1 mm using the micropositioner. Changing the height of the ink in this manner allows for precise control of the pressure applied to the ink. Displacing the reservoir upwards by 0.1 mm results in an increase in pressure of only 0.002 pounds per square inch (psi). The droplets that form while we raise and lower the reservoir of ink are separated by aperiodic intervals of time that range from 0.68 s to 0.86 s. We recorded the time that passes between consecutive droplets moving through the regions marked as "input signal," "encoded signal," and "decoded signal" in Fig. 4A. The first loop encodes the droplets by dramatically changing the sequence of intervals of time between droplets (Fig. 4C). The intervals of the encoded signal differ from the intervals of the input signal by an average of 33.3%. The second loop then decodes the signal (Fig. 4D). The intervals of the decoded signal differ from those that form the input signal by an average of 2.3%.

To encode a digital signal, we use an electronically controlled valve (*23*) to adjust the pressure applied to the reservoir of ink. We set the baseline for the code by applying 1.075 V to the valve (which applies ~3.18 psi of pressure to the reservoir of ink) and 3.26 psi of pressure to the reservoir of hexadecane. The droplets then reach the first loop separated by 0.263 ± 0.005 s. To create a value of 1, we decrease the voltage applied to the valve to 1.074 V to decrease the pressure applied to the ink and increase the intervals between droplets. To express a 0, we conversely increase the voltage applied to the valve to 1.076 V. Figure 4, E and F, shows the input, encoded, and decoded signals for the binary number 100101, which corresponds to the decimal number 37. Two bits in the input signal—the first and last 1's—do not appear in the encoded signal in Fig. 4E. In these two instances, two consecutive droplets (the intervals between which formed the peak of the signals) are separated by sufficiently long periods of time that both droplets choose the same path through the loop. Unexpectedly, the 1's are restored to the signal after the droplets pass through the second loop, although the peaks appear one interval later than in the input signal. This experiment demonstrates that the system can faithfully decode a nontrivially encoded signal.

Pine *et al.* recently studied the boundary between reversibility and irreversibility in the flow of a viscous suspension of colloidal particles. They showed that a critical number of nonlinear collisions between the particles initiates the transition to irreversibility. We used the combination of nonlinear events—the selection of paths by bubbles—and the linear nature of the viscous flow of fluids through microchannels to obtain complex behavior while preserving the reversibility of the system. We used this reversible, nonlinear system to encrypt and decrypt information. The reversible dynamics of the encoder-decoder described in Fig. 4 demonstrate that uncontrolled perturbations do not prevent the manipulation of data coded in trains of droplets. The encoder-decoder is an example of a microfluidic device that generates complex control of trajectories of droplets (e.g., directing drops into predefined branches of microfluidic networks) without using valves or switches.

References and Notes

1. F. Jousse, R. Farr, D. R. Link, M. J. Fuerstman, P. Garstecki, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **74**, 036311 (2006).
2. P. Garstecki, M. J. Fuerstman, G. M. Whitesides, *Phys. Rev. Lett.* **94**, 234502 (2005).
3. P. Garstecki, M. J. Fuerstman, G. M. Whitesides, *Nature Physics* **1**, 168 (2005).
4. H. A. Stone, A. D. Stroock, A. Ajdari, *Annu. Rev. Fluid Mech.* **36**, 381 (2004).
5. G. I. Taylor, in *National Committee for Fluid Mechanics Films* (Education Development Center, Newton, MA, 1966).
6. D. J. Pine, J. P. Gollub, J. F. Brady, A. M. Leshansky, *Nature* **438**, 997 (2005).
7. T. M. Squires, S. R. Quake, *Rev. Mod. Phys.* **77**, 977 (2005).
8. S. L. Anna, N. Bontoux, H. A. Stone, *Appl. Phys. Lett.* **82**, 364 (2003).
9. T. Thorsen, R. W. Roberts, F. H. Arnold, S. R. Quake, *Phys. Rev. Lett.* **86**, 4163 (2001).
10. T. Ward, M. Faivre, M. Abkarian, H. A. Stone, *Electrophoresis* **26**, 3716 (2005).
11. E. Cubaud, C.-M. Ho, *Phys. Fluids* **16**, 4575 (2004).

12. A. M. Ganan-Calvo, J. M. Gordillo, *Phys. Rev. Lett.* **87**, 274501 (2001).
13. P. Garstecki, H. A. Stone, G. M. Whitesides, *Phys. Rev. Lett.* **94**, 164501 (2005).
14. P. Garstecki, M. J. Fuerstman, H. A. Stone, G. M. Whitesides, *Lab Chip* **6**, 437 (2006).
15. P. Guillot, A. Colin, *Phys. Rev. E* **72**, 066301 (2005).
16. F. P. Bretherton, *J. Fluid Mech.* **10**, 166 (1961).
17. W. Engl, M. Roche, A. Colin, P. Panizza, A. Ajdari, *Phys. Rev. Lett.* **95**, 208304 (2005).
18. D. R. Link, S. L. Anna, D. A. Weitz, H. A. Stone, *Phys. Rev. Lett.* **92**, 054503 (2004).
19. P. Garstecki, M. A. Fischbach, G. M. Whitesides, *Appl. Phys. Lett.* **86**, 244108 (2005).
20. S. R. Hodges, O. E. Jensen, J. M. Rallison, *J. Fluid Mech.* **501**, 279 (2004).
21. H. Wong, C. J. Radke, S. Morris, *J. Fluid Mech.* **292**, 95 (1995).
22. Fig. S1 details the design of the microfluidic network.
23. Materials and methods are available as supporting material on *Science* Online.
24. This work was supported by the U.S. Department of Energy under award DE-FG02 00ER45852. P.G. thanks the Foundation for Polish Science for financial support. We thank the Harvard Center for Nanoscale Systems for the use of microfabrication facilities and the fast cameras.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1134514/DC1
Materials and Methods
Figs. S1 and S2
References

30 August 2006; accepted 12 December 2006
Published online 4 January 2007
10.1126/science.1134514
Include this information when citing this paper.

Microfluidic Bubble Logic

Manu Prakash* and Neil Gershenfeld

We demonstrate universal computation in an all-fluidic two-phase microfluidic system. Nonlinearity is introduced into an otherwise linear, reversible, low-Reynolds number flow via bubble-to-bubble hydrodynamic interactions. A bubble traveling in a channel represents a bit, providing us with the capability to simultaneously transport materials and perform logical control operations. We demonstrate bubble logic AND/OR/NOT gates, a toggle flip-flop, a ripple counter, timing restoration, a ring oscillator, and an electro-bubble modulator. These show the nonlinearity, gain, bistability, synchronization, cascadability, feedback, and programmability required for scalable universal computation. With increasing complexity in large-scale microfluidic processors, bubble logic provides an on-chip process control mechanism integrating chemistry and computation.

Microfluidic "lab-on-a-chip" devices, where picoliters of fluids can be precisely manipulated in microscopic channels under controlled reaction conditions, have revolutionized analytical chemistry and biosciences. Recent advances in elastomeric pneumatic microvalves (*1*) and large-scale integration (*2*) have enabled complex process control for a wide variety (*3*, *4*) of applications in single-phase microreactors. However, pneumatic elastomeric microvalves require external macroscopic solenoids for their operation, and cascadability and feedback (where a signal acts on itself) are currently lacking in microfluidic control architectures.

Several reaction chemistries have been implemented in segmented-flow two-phase microreactors, where individual nanoliter droplets traveling inside microchannels are used as reaction containers (*5*, *6*). Dielectrophoretic (*7*) and electrostatic (*8*) schemes have been proposed for on-chip droplet management, but these require external control of individual gates. Devices that exploit the dynamics of droplets inside microchannels would make high-throughput screening and combinatorial studies possible (*9*), but passive techniques (*10*, *11*) have not provided control over individual droplets.

We demonstrate bubble logic that implements universal Boolean logic in physical fluid dynamics. This provides a droplet-level, internal, inherently digital flow control mechanism for microfluidic processors. A bubble traveling in a microchannel can represent a bit of information as well as carry a chemical payload, making it possible to integrate chemistry with computation for process control. Bubble logic preserves the information representation from input to output; thus, devices can be cascaded, allowing implementation of combinational and sequential Boolean logic. A bubble can be transported to a desired location in a complex microfluidic network via a series of logic gates corresponding to an equivalent Boolean circuit.

Logic gates have been implemented chemically in chemical concentration waves in a Belousov-Zhabotinsky reaction (*12*) and in DNA (*13*). Purely hydrodynamic fluidic logic (*14*) was used to build a trajectory controller, an all-fluidic display, nondestructive memory, and a simple computer (*15*). Because the high Reynolds numbers required for inertial interactions cannot be maintained in the microscopic geometries needed for higher operating speeds and increasing integration, fluids with non-Newtonian polymer additives have been used to realize a constant flow source and a bistable gate (*16*, *17*). Boolean logic in a single-phase Newtonian fluid was implemented by changes in flow resistance (*18*), but because its input and output representation were not the same, these devices could not be cascaded. Bubble logic, based on hydrodynamic bubble-to-bubble interactions, is more similar in bit representation to theoretical billiard ball logic (*19*) based on the elastic collision of particles, and to magnetic bubble memory (*20*) relying on interactions of magnetic domains in garnet films. These schemes all conserve information, because during a logic operation a bit is neither created nor destroyed.

The pressure-driven flow of bubbles in an interconnected microfluidic network can be described with a simplified dynamic flow resistance model (*21*). Single-phase flow resistance of a channel at low Reynolds number can be approximated as $\Delta p/Q \propto \mu L/h^3w$, where $\Delta p/Q$ is defined as the hydraulic resistance per unit length, μ is the dynamic viscosity, and h and w are the height and width of the microchannel. The pressure drop due to a long bubble flowing in a channel, where the bubble radius in an unbounded fluid is greater than the channel width and the continuous phase completely wets the channel surface, is nonlinear and is proportional to $\Delta p \propto \sigma/w(3Ca^{2/3})$, where Ca is the capillary number ($Ca = \mu u/\sigma$), u is the flow velocity of the continuous phase, and σ is the surface tension between liquid and gas phase (*22*, *23*). For small flow rates, this increased flow resistance is primarily due to viscous dissipation in the thin film of liquid surrounding the bubble. With the presence of surfactant molecules on the air-water interface, viscous dissipation in the lubrication film further increases as a result of the no-slip boundary conditions at the interface. In this case, the pressure drop across a finite-length bubble is also linearly dependent on the bubble length until it reaches a critical value, beyond which it is constant (*24*). When a bubble traveling in a microchannel arrives at a bifurcation with low capillary number (where the bubble does not split because surface tension dominates the viscous stress), it chooses the branch with highest instantaneous flow (*25*, *26*).

With an increased flow resistance due to the presence of a bubble in a microchannel, flow lines in surrounding interconnected channels can be perturbed. The nonlinearity in such a system arises from the introduction of interfacial force terms from the boundary conditions due to the presence of a free surface at the fluid interfaces (*27*). These nonlinear time-dependent interactions are the basis of our bubble logic gates. In the implementation described here, we used water as the liquid medium [with added surfactant 2% (w/w) Tween 20 to stabilize the interfaces] and nitrogen bubbles. Planar bubble logic devices were fabricated in poly(dimethyl siloxane) (PDMS) by single-layer soft lithography and plasma bonding to Pyrex substrates.

Center for Bits and Atoms, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: manup@mit.edu

Figure 1 (top row; see also movie S1) shows an AND-OR bubble logic gate that evaluates both AND (·) and OR (+) simultaneously, as is necessary to satisfy bit conservation. As shown in the center column, the first bubble arriving at the junction always enters A·B (the wider channel, with less resistance), increasing the output flow resistance of A·B and thus directing a bubble arriving later to A+B. The time trace plotted for all four channels shows that the two bubbles interact only if they arrive within a window τ_0 (for this gate τ_0 ~ 0.5 ms at flow rate Q = 0.25 μL/s) determined by the residence time of the bubble in the gate geometry.

Fan-out for the output signal from one gate to act as an input signal for multiple gates can be implemented by splitting bubbles at a T junction (*25*), requiring restoration of the bubble size. Figure 1 (bottom row; see also movie S2) demonstrates a universal A·B gate, which implements a NOT and AND with gain so that a smaller bubble can switch a larger one. There are two counteracting asymmetries: an input channel with an asymmetric T junction (bottom), and a narrow stream of injected flow from the control channel (top) into the wider of the two bifurcations. By introducing a bubble into the control channel, injected side flow can be dynamically turned on and off, thereby enabling control of the direction of flow of the output bubble arriving at the bifurcation (bottom row, center column). The change in injected flow from the control channel (ΔQ) when a bubble passes through is nonlinearly related to the size of the bubble, providing gain as plotted in Fig. 1 (bottom row, right column) against the dimensionless bubble size (bubble length/channel width).

Segmented-flow reactors often operate at kilohertz frequencies, where the limiting factor

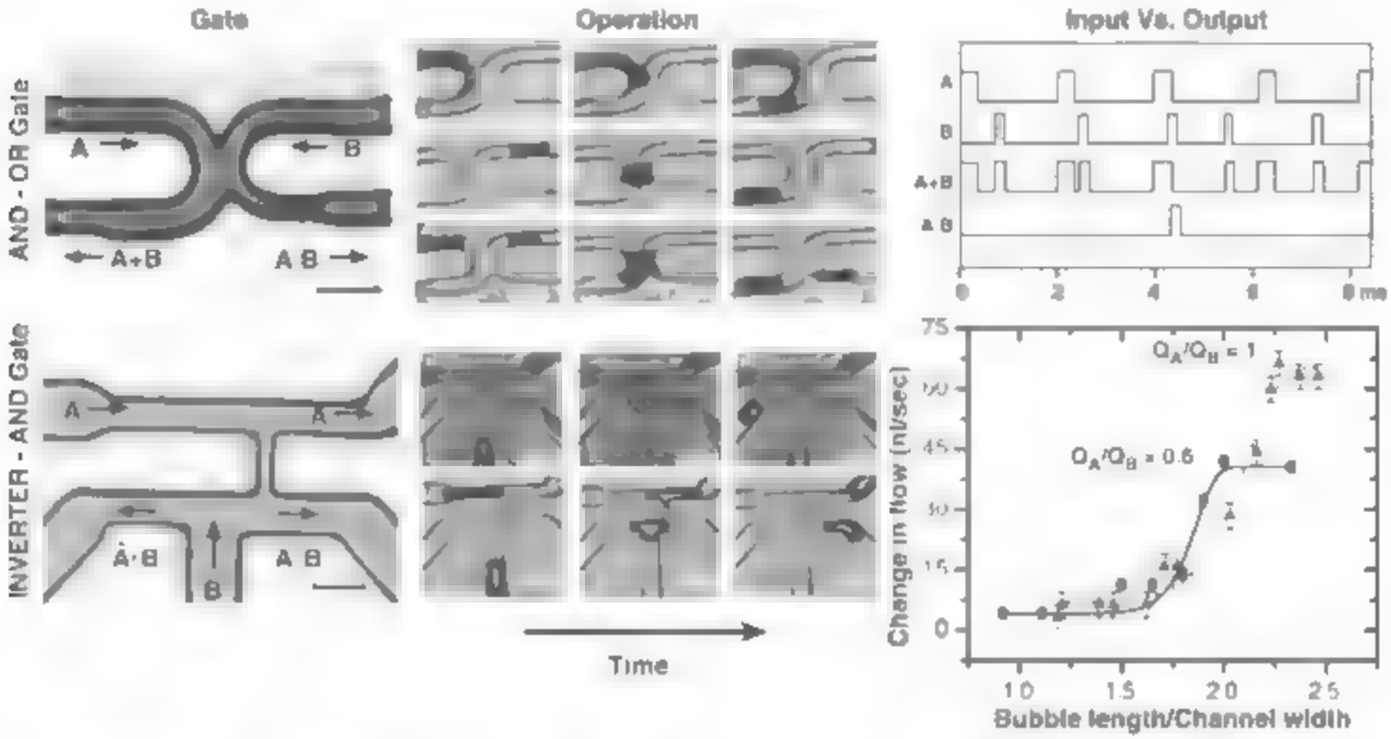


Fig. 1. Universal, microfluidic bubble logic. The top row shows a two-input AND-OR gate. The channel height is 70 μm; scale bar, 100 μm. With a water flow rate of 0.25 μL/s and nitrogen bubbles with a pressure of 0.5 psi, the gate propagation delay is 2 ms. The bottom row depicts a universal Ā·B gate with gain that can be used to switch a larger bubble by a smaller one. For the bubbles shown, the ratio of the size of the input to the control is 1.2, and the graph shows the nonlinear dependence of the inlet flow on the control bubble size for two ratios of the control to the inlet flow rate.

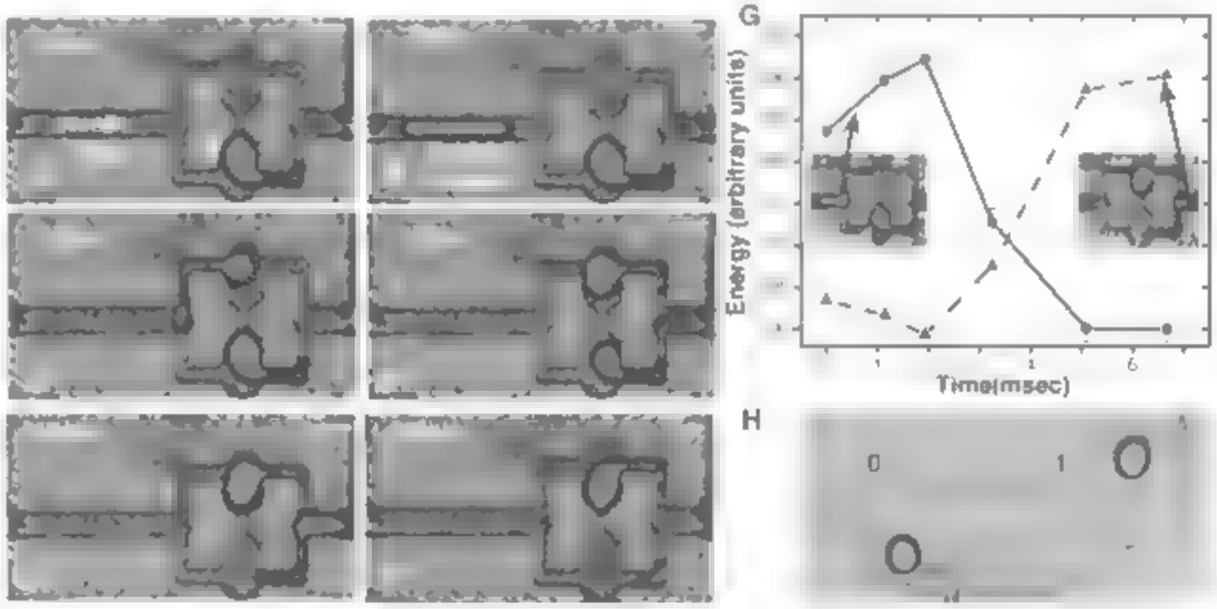


Fig. 2. Bistability. (**A** to **F**) Toggling of a flip-flop memory; with a water flow rate of 0.25 μL/s, the switching time is 8 ms. The channel height is 70 μm; scale bar, 100 μm. (**G**) Change in free surface energy (*31*) for the toggle (solid line) and stored (dashed line) bubbles during switching. (**H**) Flip-flops cascaded to form a ripple counter.

for high-throughput screening is the rate of information extraction from individual droplets. We present a bistable mechanism (fig. S1 and movie S3) capable of on-demand trapping and release of individual bubbles, implemented as a flip-flop memory. A bubble minimizes its surface energy by adopting a shape with the smallest surface area. The flip-flop geometry (Fig. 2A) presents an incoming bubble with two elliptical lobes when the surface energy of the bubble is at a minimum, as shown in the plot of energy versus time (Fig. 2G and fig. S1A). The device holds a single bubble indefinitely until it is toggled by another bubble arriving at the inlet, dislodging the stored bubble by flow through the channel connecting the lobes. Although the incoming bubbles are much longer than the Rayleigh-Plateau criteria for breakup (*25*), the presence of a bubble in the flip-flop ensures that the incoming bubble travels to a single lobe without breakup (the bifurcation diagram is plotted in fig. S1B and the repeatability in the time trace in fig. S1C). We have used this mechanism to implement a bistable flow switch (fig. S2), with a switching time an order of magnitude less than that for commercial macroscopic elements such as solenoid valves (*7*), as well as a cascaded ripple counter (Fig. 2H).

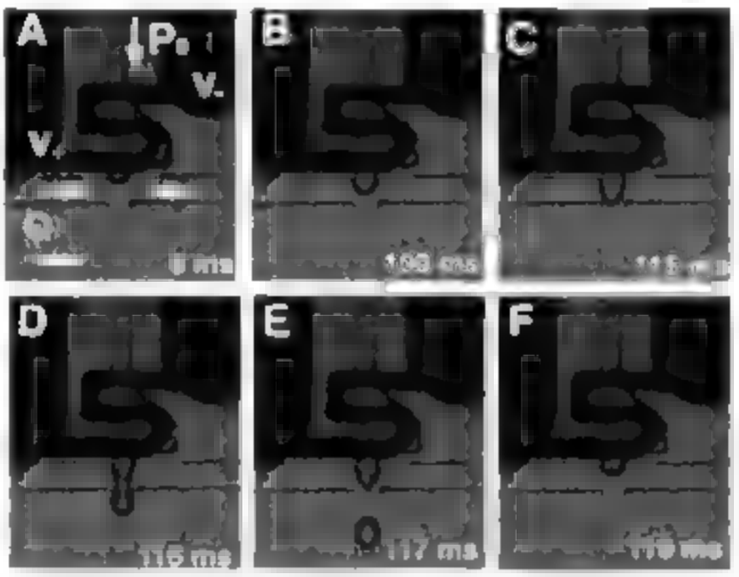


Fig. 3. Programmable bubble generator. A platinum microheater (width 50 μm, thickness 200 nm, R = 95 ohms) followed by a 2-μm silicon dioxide dielectric barrier is embedded below a planar flow-focusing geometry (channel height 70 μm, water flow rate 0.83 μl/s, nitrogen pressure 5 psi). Scale bar, 100 μm. (**A** to **C**) A 21-V dc pulse is applied to the heater for 100 ms, resulting in the growth of a meniscus. Arrows indicate direction of flow. (**D**) A gas filament penetrates the water. (**E**) The filament breaks into a single drop via flow focusing (*28*), followed by (**F**) the interface retracting.

To provide an electronic interface to bubble logic devices, we developed a thermal electro-bubble modulator (Fig. 3 and movie S4) capable of generating bubbles on demand synchronized to an electric pulse. Methods for high-frequency continuous production of monodisperse microbubbles and droplets in microfluidic devices have been extensively studied (*27*–*29*). Electrogeneration of on-demand single aqueous droplets (*30*) requires high on-chip electric fields on the order of ~1 kV. Our thermal electro-bubble generator uses an integrated microheater and modified flow-focusing geometry (Fig. 3), enabling operation at a much lower voltage (21 V). For the case of pressure-driven flow (Fig. 3A), there is a static balance at the air-water interface $\Delta P + \tau_v \sim \sigma k$ (*29*), where ΔP is the difference in pressure, τ_v is the viscous stress, and k represents the mean curvature. An applied temperature pulse reduces the surface tension σ at the air-water interface, allowing a gas filament to penetrate the liquid (Fig. 3D), which breaks to form a single bubble (Fig. 3F).

A standard test of cascadability and feedback in a new logic family is steady-state operation of a ring oscillator. We implemented this with three identical AND gates connected via three delay lines in a ring structure (Fig. 4A and movie S5) with constant-frequency T-junction bubble generators at the inputs. A bubble propagating around the ring delay line (Fig. 4A, arrow) increases the resistance of the outgoing channel when it arrives at the input of one of the three AND gates, generating a pressure pulse that launches another bubble in response. The oscillation frequency of this device can be written as $f \propto 1/(3l/v + \tau_d)$, where f is the oscillator frequency, l is the length of the delay line, v is the mean velocity of the bubble traveling in the delay line, and τ_d is the propagation delay of the AND gate. This frequency can be tuned by increasing the flow rate of the continuous phase (Fig. 4B).

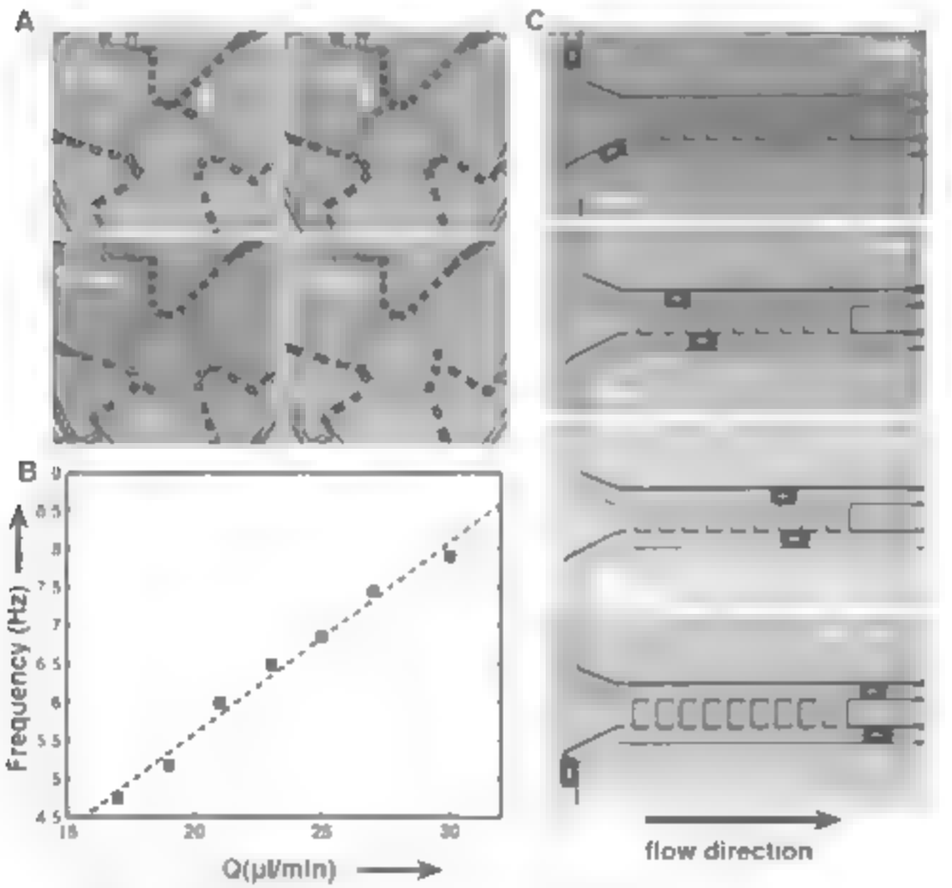


Fig. 4. Feedback, cascadability, and synchronization. (**A**) Three AND gates connected as a ring oscillator with constant-frequency bubble generators at the inputs. A bubble arriving at the input delay line releases one at the output delay line. Scale bar, 100 μm. (**B**) Dependence of the oscillation frequency on the flow rate Q. (**C**) A fluidic ladder network for timing restoration. With a flow rate of 0.5 μl/s, a restoration of ~10 ms is achieved over a span of ~40 ms. Scale bar, 100 μm.

The operation of bubble logic requires the relative arrival of bubbles at a gate to be within a transit time. Scalability requires restoration of relative timing errors. We achieved this via a planar fluidic resistance ladder network. This geometry (Fig. 4C and movie S6) places inter-

connecting fluid channels (continuous-phase flow resistance r) between two data-carrying channels (continuous-phase flow resistance R with $r > R$). A single bubble traversing the ladder is slowed down by flow through the alternate path. When both bubbles are present simultaneously, there is a net flow from the channel with the leading bubble to the one with the lagging bubble, creating a relative velocity gradient until the bubbles are synchronized (fig. S3).

Because bubble logic chips have no moving parts, they can be fabricated in a wide variety of materials, including silicon and glass, that are compatible with reaction chemistries unsuitable for PDMS channels. Moreover, because they operate at low Reynolds and capillary numbers, further reduction in size is feasible with faster switching times. The device mechanisms do not depend on non-Newtonian fluid properties; hence, matching dimensionless flow parameters will allow bubble logic circuits to be designed using different fluids, such as water droplets in oil or oil droplets in water.

The universal logic gates, toggle flip-flop, ripple counter, synchronizer, ring oscillator, and electro-bubble modulator presented here exhibit nonlinearity, bistability, gain, synchronization, cascading, feedback, and programmability. Having shown the required properties of a scalable logic family, they can be used to create complex microfluidic circuits capable of performing arbitrary fluid process control and computation in an integrated fashion. Such circuits may reduce the size, cost, and complexity of current microfluidic systems, thereby enabling the development of very-large-scale microfluidic reactors for use in areas including combinatorial chemistry and drug discovery. These bubble logic processors, where a bit of information can also carry a chemical payload, merge chemistry with computation.

References and Notes

1. M. A. Unger, H.-P. Chou, T. Thorsen, A. Scherer, S. R. Quake, *Science* **288**, 113 (2000).
2. T. Thorsen, S. J. Maerkl, S. Quake, *Science* **298**, 580 (2002); published online 26 September 2002 (10.1126/science.1076996).
3. C. C. Lee *et al.*, *Science* **310**, 1793 (2005).
4. F. K. Balagaddé, L. You, C. L. Hansen, F. H. Arnold, S. R. Quake, *Science* **309**, 137 (2005).
5. K. Jensen, A. Lee, *Lab Chip* **4**, 31N (2004).
6. B. Zheng, L. S. Roach, R. F. Ismagilov, *J. Am. Chem. Soc.* **125**, 11170 (2003).
7. P. R. C. Gascoyne *et al.*, *Lab Chip* **4**, 299 (2004).
8. D. Link *et al.*, *Angew. Chem. Int. Ed.* **45**, 2556 (2006).
9. M. Joanicot, A. Ajdari, *Science* **309**, 887 (2005).
10. Y.-C. Tan, J. S. Fisher, A. I. Lee, V. Cristini, A. P. Lee, *Lab Chip* **4**, 292 (2004).
11. G. Cristobal, J.-P. Benoit, M. Joanicot, A. Ajdari, *Appl. Phys. Lett.* **89**, 034104 (2006).
12. O. Steinbock, P. Kettunen, K. Showalter, *J. Phys. Chem.* **100**, 18970 (1996).
13. G. Seelig, D. Soloveichik, D. Y. Zhang, E. Winfree, *Science* **314**, 1585 (2006).
14. A. Conway, *A Guide to Fluidics* (Macdonald, London, 1972).
15. C. A. Belsterling, *Fluidic Systems Design* (Wiley Interscience, New York, 1971).
16. A. Groisman, M. Enzelberger, S. R. Quake, *Science* **300**, 955 (2003).
17. A. Groisman, S. R. Quake, *Phys. Rev. Lett.* **92**, 094501 (2004).
18. T. Vestad, D. W. Marr, T. Munakata, *Appl. Phys. Lett.* **84**, 5074 (2004).
19. E. Fredkin, T. Toffoli, *Int. J. Theor. Phys.* **21**, 219 (1982).
20. H. Chang, Ed., *Magnetic Bubble Technology: Integrated-Circuit Magnetics for Digital Storage and Processing* (IEEE Press, New York, 1975).
21. F. Jousse, G. Lian, R. Janes, J. Melrose, *Lab Chip* **5**, 646 (2005).
22. F. P. Bretherton, *J. Fluid Mech.* **10**, 166 (1961).
23. H. Wong, C. Radke, S. Morris, *J. Fluid Mech.* **292**, 95 (1995).
24. C. W. Park, *Phys. Fluids* **4**, 2335 (1992).
25. D. Link, S. Anna, D. Weitz, H. Stone, *Phys. Rev. Lett.* **92**, 054503 (2004).
26. M. J. Fuerstman, P. Garstecki, G. M. Whitesides, *Science* **315**, 828 (2007); published online 4 January 2007 (10.1126/science.1134514).
27. T. Thorsen, R. W. Roberts, F. H. Arnold, S. R. Quake, *Phys. Rev. Lett.* **86**, 4163 (2001).
28. P. Garstecki *et al.*, *Appl. Phys. Lett.* **85**, 2649 (2004).
29. T. Ward, M. Faivre, M. Abkarian, H. Stone, *Electrophoresis* **26**, 3716 (2005).
30. M. He, J. S. Kuo, D. T. Chiu, *Appl. Phys. Lett.* **87**, 031916 (2005).
31. Surface energy for a bubble trapped in a channel was evaluated from photomicrographs captured by a high-speed camera (Phantom V5.0, Vision research; minimum exposure time 10 μs). The grayscale images from the camera were digitized, a threshold filter was applied, and an edge detector was used to locate the bubbles and find their perimeter.

32. Supported by the MIT Center for Bits and Atoms (NSF grant CCR-0122419).

Supporting Online Material
www.sciencemag.org/cgi/content/full/315/5813/832/DC1
Figs. S1 and S2
Movies S1 to S6

30 October 2006; accepted 5 January 2007
10.1126/science.1136907

Chemical and Spectroscopic Evidence for an Fe^{V}-Oxo Complex

Filipe Tiago de Oliveira,[1*] Arani Chanda,[1*] Deboshri Banerjee,[1] Xiaopeng Shan,[2] Sujit Mondal,[1] Lawrence Que Jr.,[2†] Emile L. Bominaar,[1†] Eckard Münck,[1†] Terrence J. Collins[1†]

Iron(V)-oxo species have been proposed as key reactive intermediates in the catalysis of oxygen-activating enzymes and synthetic catalysts. Here, we report the synthesis of $[Fe(TAML)(O)]^-$ in nearly quantitative yield, where TAML is a macrocyclic tetraamide ligand. Mass spectrometry, Mössbauer, electron paramagnetic resonance, and x-ray absorption spectroscopies, as well as reactivity studies and density functional theory calculations, show that this long-lived (hours at −60°C) intermediate is a spin $S = 1/2$ iron(V)-oxo complex. Iron TAML systems have proven to be efficient catalysts in the decomposition of numerous pollutants by hydrogen peroxide, and the species we characterized is a likely reactive intermediate in these reactions.

High-valent iron-oxo intermediates are used in biological systems to carry out challenging oxidations in many scenarios (*1–3*). It has long been known that heme-based systems, biological as well as synthetic, do not attain the iron(V) state. Rather, the so-called "Compound I" intermediate involves the Fe^{IV}(oxo)(porphyrin-radical-cation), $Fe^{IV}(O)(P^{+•})$, which is well established in peroxidases and catalases and presumed with considerable justification to be the reactive oxidizing species of the cytochrome P450s (*1*). Recently, Newcomb and co-workers reported laser flash photolysis experiments of an iron-corrole complex (*4*) and cytochrome P450 mutant CYP 119 (*5*), in which optically detected transients were tentatively assigned to Fe^{V}-oxo species. Que and co-workers, on the basis of ^{18}O labeling experiments, have postulated that an $HO\text{-}Fe^{V}$-oxo oxidant is an intermediate in the large family of Rieske dioxygenase enzymes (*3*). Iron(V) complexes are exceedingly rare, but recently, Wieghardt and co-workers reported spectroscopic evidence for a non-heme $S = 1/2$ iron(V)-nitrido complex produced by photolysis of an azidoiron(III) precursor (*6*).

Tetraamido macrocyclic ligands (TAMLs) can stabilize a variety of high-valent iron complexes, including a high-spin ($S = 2$) iron(IV) complex, intermediate-spin ($S = 1$) iron(IV) complexes, oxo-bridged diiron(IV) dimers, and an $S = 1$ iron(III)(TAML-radical-cation) complex [all reviewed in (*7*)]. The structurally characterized oxo-bridged complex, $[(Fe^{IV}B^{*})_2(\mu\text{-oxo})]^{2-}$, **2**, ($B^{*}$, a TAML, is shown in Fig. 1) was obtained by reacting the monomeric iron(III) precursor, $[Fe^{III}B^{*}(H_2O)]^-$, **1**, with air (*8*). Fe-TAML activation of peroxide has been shown to be useful in numerous applications (*9*, *10*). The deprotonated TAML (a tetraanion) presents the iron with four exceptionally strong amido-N σ-donors. Therefore, it has been reasonable to expect (*11*) that this macrocyclic ligand would be capable of stabilizing an oxo-iron(V) complex when an Fe^{III}(TAML) complex is treated with an oxygen-atom transfer agent such as peroxide. A variety of iron(IV) complexes have been synthesized by reacting iron(III) precursors with oxygen donors

[1]Department of Chemistry, Carnegie Mellon University, Pittsburgh, PA 15213, USA. [2]Department of Chemistry and Center for Metals in Biocatalysis, University of Minnesota, Minneapolis, MN 55455, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: que@chem.umn.edu (L.Q.); eb7g@andrew.cmu.edu (E.L.B.); emunck@cmu.edu (E.M.); tc1u@andrew.cmu.edu (T.J.C.)

in organic solvents such that the reaction can be monitored optically at temperatures as low as −80°C (*12*, *13*). A similar approach was applied here to Fe^{III} precursors to monitor short-lived intermediates in the reactions of TAML complexes. We report the synthesis of a d^3 iron(V)-oxo complex, in nearly quantitative yield, and its characterization with optical, electron paramagnetic resonance (EPR), Mössbauer and x-ray absorption spectroscopies, as well as electrospray ionization mass spectrometry (ESI-MS), reactivity studies, and density functional theory (DFT).

The reaction of $[PPh_4][Fe^{III}B^*(H_2O)]$ (**1**, 1 to 2 mM) with 2 to 5 equivalents (equiv) of *m*-chloroperbenzoic acid (*m*CPBA) at −60°C in *n*-butyronitrile (*14*) produces within about 10 s the diiron(IV) dimer **2** (Fig. 1, dotted line). The reaction then proceeds through an as-yet uncharacterized EPR-silent Fe(IV) intermediate (a Mössbauer spectrum is shown in fig. S3) to yield within 15 min after the addition of *m*CPBA, a deep green species, **3**. Complex **3** (Fig. 1, solid line) exhibits distinct absorption maxima at 445 nm (where the extinction coefficient, ε, is 5400 M^{-1} cm^{-1}, calibrated by reference to Mössbauer spectra) and 630 nm (ε = 4200 M^{-1} cm^{-1}). At −60°C, **3** decays by 10% in 90 min, but it is stable for at least one month at 77 K. Complex **3** was also obtained in acetonitrile at −40°C, but at this temperature it decays by 10% in about 5 min, probably by attacking the solvent (*11*). In the course of our studies, we learned that **3** can be obtained quantitatively with 1 equiv of *m*CPBA when small amounts of potential ligands such as water, pyridine, and benzoic acid were present. The spectral features of **3** were independent of the presence of these potential ligands.

ESI-MS examination of *n*-butyronitrile solutions of **3** revealed one prominent ion at a mass-to-charge ratio (*m*/*z*) of 442.2, the mass and isotopic distribution pattern of which correspond to a formulation of $[Fe^{V}B^*(O)]^+$ [see supporting online material (SOM) text and fig. S1]. Only one other peak was observed, at *m*/*z* = 581.1, and it corresponded to a decay product of **3**, namely $[Fe^{IV}B^*(m\text{-chlorobenzoate})]^+$. Introduction of $H_2{}^{18}O$ into the solvent media yielded up to 35% of $[Fe^{V}B^*(^{18}O)]^+$ (*m*/*z* = 444.2) after 30 min of incubation. Presently, we do not know at what stage of the reaction the ^{18}O exchange occurs.

In the 28 K EPR spectrum of **3** (Fig. 2A, solid line), prepared at −60°C in the glass-forming solvent, *n*-butyronitrile, the major feature belongs to an $S = 1/2$ species with *g* values of about 1.99, 1.97, and 1.74 [*g* values are the eigenvalues of tensor **g**, which couples the electronic spin, **S**, to the applied magnetic field, **B** (Eq. 1)]. The dashed line represents a spectral simulation drawn to correspond to an $S = 1/2$ spin concentration of 0.92(5) (estimated error from spectral simulation in the least significant digit) spins per Fe molecule. A minor species, present in all samples, with $g_1 = g_2 = 2.06$ and $g_3 = 2.00$ to 2.02 contributes about 3% of the total spin concentration.

We studied samples of **3** with Mössbauer spectroscopy between 4.2 and 140 K in applied magnetic fields up to 8.0 T. Three representative spectra of a sample prepared at −60°C are shown in Fig. 2B; additional spectra are displayed in fig. S2. The majority species contributes nearly 95% of the absorption, and thus we associate it with the majority $S = 1/2$ species observed by EPR. The Mössbauer spectra of **3** were simulated with the $S = 1/2$ spin Hamiltonian

$$H = \beta \mathbf{S} \cdot \mathbf{g} \cdot \mathbf{B} \quad \mathbf{S} \cdot \mathbf{A} \cdot \mathbf{I} \quad g_N \beta_N \mathbf{B} \cdot \mathbf{I}$$

$$e^2 Q V_{zz} / 12 [3I_z^2 - 15/4 + \eta (I_x^2 - I_y^2)] \quad (1)$$

In Eq. 1, β is the Bohr magneton; **S** is the electron spin operator; **g** is the *g* tensor; **B** is the applied magnetic field; **A** is the ^{57}Fe magnetic hyperfine tensor; **I** is the nuclear spin operator; $g_N\beta_N$ is the nuclear gyromagnetic ratio; I_x, I_y, and I_z are the Cartesian components of **I**; and the last term describes the interaction between the nuclear quadrupole moment *Q* and the electric field gradient (EFG) tensor (principal components V_{xx}, V_{yy}, V_{zz}; $\Delta E_Q = (eQV_{zz}/2)\sqrt{1 + \eta^2/3}$ with $\eta = (V_{xx} - V_{yy})/V_{zz}$).

At 4.2 K, the electronic spin of **3** relaxes slowly on the time scale of Mössbauer spectroscopy (about 10^{-7} s), and thus we observed paramagnetic hyperfine structure. At 140 K, the spin system approaches the fast relaxation limit, and the magnetic hyperfine interactions were essentially averaged out. The spectrum 1 of Fig. 2B exhibits a doublet (solid line) with quadrupole splitting $\Delta E_Q = 4.25$ mm/s and isomer shift $\delta = 0.46(2)$ mm/s (relative to Fe metal at 298 K). $\delta = -0.42(3)$ mm/s at 4.2 K. This δ value is considerably lower than the lowest δ observed for any Fe-TAML complex, specifically $\delta = -0.18$ mm/s (at 4.2 K) for an $S = 1$ iron(IV) complex (table S5). Because **3** has a half-integral electronic spin, we assigned this species to a low-spin d^3 $[Fe^{V}B^*(O)]^+$ complex, in accord with the mass spectroscopic data. The solid lines in Fig. 2B are spectral simulations based on Eq. 1 according to the parameters listed in Table 1. The theoretical curves in i and iii in Fig. 2B represent 95% of the total Fe in the sample, in excellent agreement with the EPR data [0.92(5) spins per Fe molecule].

Structural evidence for the nature of **3** derives from x-ray absorption spectroscopic studies at the Fe K edge. (The Fe K edge corresponds to the transition of the iron 1*s* electron to the continuum of electron states.) A sample in frozen CH_3CN solvent, 70% pure in **3**, exhibits a K edge at 7128.3 eV, an energy value 1.4 eV higher than that associated with its iron(III) precursor **1** and thus consistent with the higher oxidation state of **3**. The 1*s* → 3*d* pre-edge transition found for the sample of **3** is quite intense, with an area of 48 units compared with 17 units for **1** (Fig. 3A). Taking into account that only 70% of the sample is **3** (the remainder being Fe(IV)], a pre-edge area of at least 60 units can be estimated for

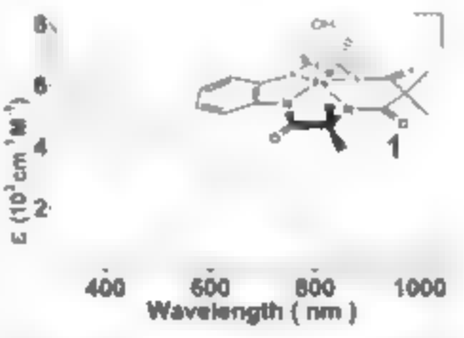


Fig. 1. Electronic absorption spectra of **1** (dashed line) and of **2** (dotted line) and **3** (solid line) as they formed during a reaction in *n*-butyronitrile at −60°C by the addition of 3 equiv of *m*CPBA. The inset shows the iron(III)-aqua complex **1**, $[Fe^{III}B^*(H_2O)]^-$, where the free-base H_4B^* tetraamide is 3,3,6,6,9,9-hexamethyl-3,4,8,9-tetrahydro-1*H*-1,4,8,11-benzotetraazacyclotridecine-2,5,7,10(6*H*,11*H*)-tetraone.

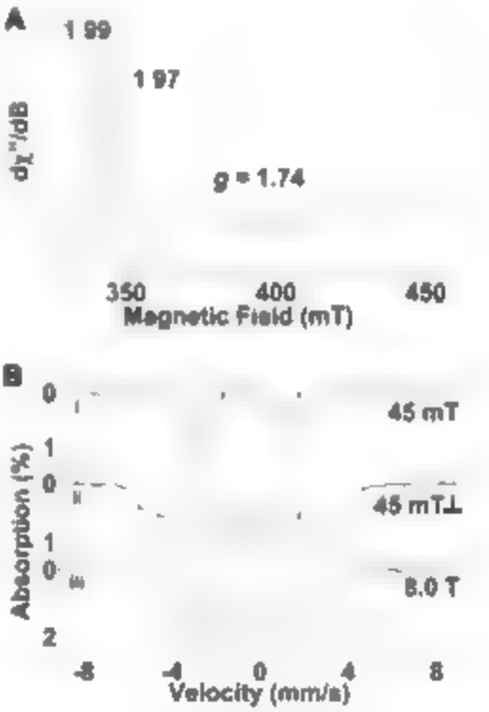


Fig. 2. (**A**) X-band EPR spectrum of about 2 mM ^{57}Fe-enriched **3** in *n*-butyronitrile. Conditions: T = 28 K; frequency, 9.66 GHz; microwave power, 0.02 mW; modulation, 1 mT. The dashed line is a spectral simulation for the majority species, **3**, using the following parameters: $g_x = 1.985$, $g_y = 1.966$, $g_z = 1.735$, with *g* values assumed to have a Gaussian distribution with $\sigma_{g_x} = 0.02$, $\sigma_{g_y} = 0.03$, and $\sigma_{g_z} = 0.05$, using a packet width of 1 mT. The *y* axis of the EPR spectrum is expressed in arbitrary units. (**B**) Mössbauer spectra of about 2 mM ^{57}Fe-enriched $[Fe^{V}B^*(O)]^+$, **3**, in *n*-butyronitrile recorded at 140 K (i) and 4.2 K (ii and iii) in magnetic fields indicated; incident γ beam perpendicular (ii) and parallel (i and iii) to the applied field. The solid lines are spectral simulations based on Eq. 1 using the parameters of Table 1, as drawn they represent 95% of the total absorption. Arrows (in i and ii) indicate a decay product (3% of Fe) of **3** with ΔE_Q = 3.4 mm/s and δ = 0.08 mm/s.

pure **3**. This value is distinctly larger than those observed for known $S = 1$ six-coordinate oxo-iron(IV) species (25 to 38 units) (*12*).

Direct evidence for an iron-oxo unit is obtained from an analysis of the extended x-ray absorption fine structure (EXAFS) region of the x-ray absorption spectrum (Fig. 3B). The best fit obtained, considers three shells: 0.7 O scatterer (fixed on the basis of the Mössbauer analysis of this sample) at 1.58 Å, four N/O scatterers at 1.87 Å, and five C scatterers at 2.82 Å (see SOM text for details of the fitting protocol). Whereas the scatterers at 1.87 and 2.82 Å are consistent with low-*Z* atoms from the macrocyclic ligand as found in the EXAFS fits for **1**, the 1.58 Å scatterer is unique for **3** and assigned to the oxo atom of an Fe-O unit. Thus, **3** can be formulated as $[Fe^{V}B^{*}(O)]^{+}$ by a combination of evidence from electrospray mass spectrometry, EPR, Mössbauer analysis, and x-ray absorption spectroscopy.

Consistent with the above formulation, **3** displays the signature reactivity features of a strongly oxidizing iron-oxo complex. Thus, reaction of excess triphenylphosphane (10 equiv) with preformed **3** (1 mM in 500:1 volume by volume acetonitrile to water, with the use of 1 equiv of *m*CPBA at −40°C) quantitatively produced triphenylphosphane oxide (1 mM $OPPh_3$), regenerating **1**. When this reaction was carried out in the presence of $H_2{}^{18}O$ (0.2% by volume), incorporation of ^{18}O in the oxidized product was observed by gas chromatography mass spectrometry studies. As described in the SOM text (table S3), oxidation proceeded readily with other substrates under similar conditions, also showing ^{18}O incorporation when $H_2{}^{18}O$ was present. Thioanisole was oxidized quantitatively to the corresponding sulfoxide. Substoichiometric reactions with styrene and cyclooctene produced the corresponding epoxides, reaction with ethylbenzene gave a mixture of 1-phenylethanol and acetophenone. In these cases where **3** did not revert completely to iron(III), a mixture of lower-valent iron complexes was produced, including the diiron(IV) dimer, **2**. Oxidation of 9,10-dihydroanthracene proceeded quantitatively to anthracene and **2**.

DFT calculations using Gaussian with the functional B3LYP and basis set 6-311G clearly favor the low-spin ($S = 1/2$) $[Fe^{V}B^{*}(O)]^{+}$ configuration as the ground state of **3**. A geometry-optimized structure together with an orbital diagram is shown in Fig. 4. The calculated Fe-O bond length of 1.60 Å is in excellent agreement with our EXAFS results. This distance is 0.04 to 0.1 Å shorter than those reported for Fe^{IV}=O complexes (*13*). The contraction of the Fe-O bond reflects an increase in the formal bond order from 2 in Fe^{IV}=O to 2.5. The Fe atom is displaced out of the 4-N plane by 0.5 Å. DFT calculations showed the lowest excited $S = 3/2$ state to be an Fe^{IV}-TAML cation radical state $[(Fe^{IV}(B^{*})^{\bullet+}(O)]^{+}$, at 1900 cm^{-1} rather than the metal-based state obtained by a spin flip transition from xy to xz (*15*).

One-electron oxidations of iron-oxo-porphyrinates are generally ligand based rather than metal based as in **3** (*16*). A major factor contributing to this difference is the stronger σ-donor capacity of the deprotonated amide nitrogens of a TAML tetraanion compared with the pyrroles of the porphyrin dianion. Hence, the TAML injects more electronic charge into the 3*d* shell than does a porphyrin, thereby raising the energy of the redox-active metal orbital and protecting the iron(V) against electron transfer from the phenyl moiety.

The values calculated for δ, ΔE_Q, η, and $A_{x,y,z}$ (Table 1) provide strong support to our iron(V) assignment. The isomeric shift calculated for $[Fe^{V}B^{*}(O)]^{+}$, δ = −0.39 mm/s, is in excellent agreement with the data and is substantially lower than that calculated for $[Fe^{IV}(B^{*})^{\bullet+}(O)]^{+}$, δ = −0.15 mm/s (calculated δ values for Fe-TAML complexes agree within 0.02 to 0.03 mm/s with the experimental values; see table S5). The removal of a 3*d* electron (reducing the shielding of the nuclear charge) and the shortening of the Fe-O bond distance (enhancing the covalent 4*s* population) each increase the electron density at

Table 1. Mössbauer, EPR, and DFT parameters of $[Fe^{V}B^{*}(O)]^{+}$. Numbers in parentheses indicate the estimated errors in the least significant digits. Because DFT calculations generally produce contact terms that are too small (isotropic), the *A* values were obtained by adding the calculated spin-dipolar contribution to the experimental $A_{iso} = (A_x + A_y + A_z)/3$ and then adding a correction for the orbital contribution as described in the SOM text.

	g_x, g_y, g_z	$(A_x, A_y, A_z)/g_N\beta_N$(T)	ΔE_Q(mm/s)	η	δ(mm/s)
Exp	1.99, 1.97, 1.74	−49.3(10), −1.5(15), −16.3(15)	+4.25(10)	0.65(10)	−0.42(3)*
DFT	–	43.6, 9.5, 15.0	+4.51	0.72	−0.39

*At 4.2 K relative to Fe metal at 298 K.

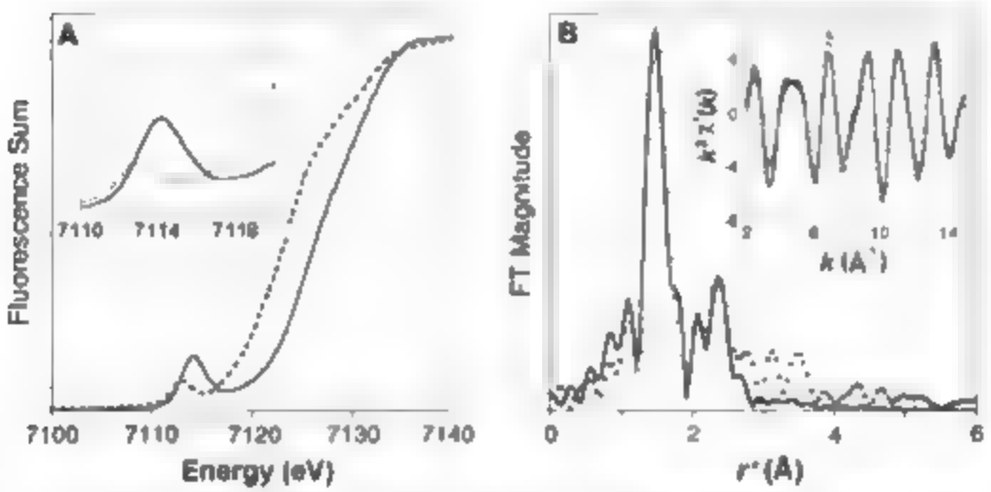


Fig. 3. (**A**) X-ray absorption near-edge spectra (Fe K edge) of **3** (solid line) and **1** (dashed line). The inset shows the magnified 1*s* → 3*d* transition features. (**B**) Fourier transforms of the Fe K edge EXAFS data ($k^3\chi(k)$) (dots) and the fit (solid line) for **3** with 0.7 O scatterer at 1.58 Å, four N/O scatterers at 1.87 Å, and five C scatterers at 2.82 Å. The inset shows Fourier-filtered EXAFS spectra ($k^3\chi$ vs. k) (dots) and the fit (solid line). Fourier transform range k = 2 to 14.9 $Å^{-1}$; back-transform range: r' = 0.70 to 3.00 Å. The fitting protocol is detailed in SOM text.

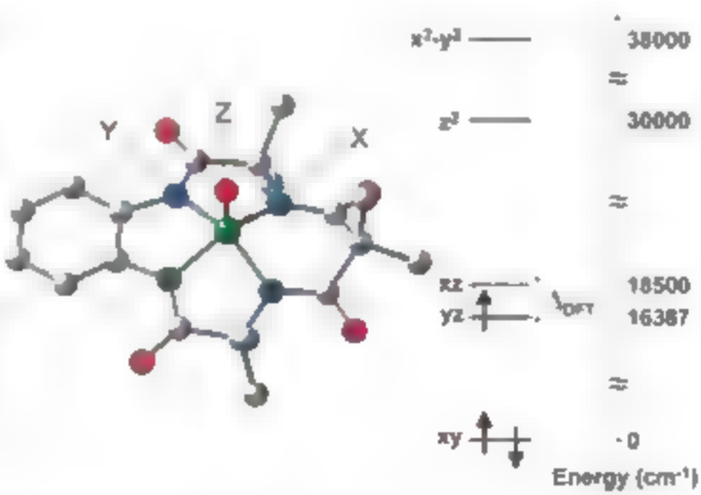


Fig. 4. Geometry-optimized structure of **3** as obtained from DFT calculations (green, iron; red, oxygen; blue, nitrogen; and gray, carbon). For clarity, hydrogen atoms are not shown. Electronic d-orbital diagram of the Fe^{V} center is shown, as derived from DFT. Selected bond lengths: Fe–O, 1.60 Å; Fe–N, 1.86 Å (two nitrogens along *x* direction) and 1.91 Å (two nitrogens along *y* direction); the Fe is 0.5 Å above the plane defined by the four amide nitrogens.

the Fe nucleus, thus decreasing δ. In contrast, the removal of an electron from the redox-active ligand orbital has a minor effect on δ (e.g., δ = −0.15 mm/s for $[Fe^{IV}(B^*)^{+}(O)]^-$ compared with δ = −0.12 mm/s for $[Fe^{IV}(B^*)(O)]^{2-}$) (table S5). Inspection of the DFT results shows that the Fe-O bond of **3** has nearly the same covalency as the Fe-O bonds in Fe^{IV}=O complexes (*12*, *17*, *18*). The calculated unpaired spin density on the oxo group of **3** is 0.22, which is less than expected (≈0.4) by comparison with Fe^{IV}=O complexes (≈0.8); the difference can be traced to a spin polarization of the bonding (p_x − d_{xz}) orbital. The exchange potential generated by the d^{α}_{yz} electron places the virtual d^{α}_{xz} level below the d^{β}_{xz} level. As a consequence, the transfer of electron density from p^{α}_x to d^{α}_{xz} is greater than from p^{β}_x to d^{β}_{xz}, resulting in a negative spin density in p_x, which, combined with the positive spin density for p_y (~0.4), yields a net spin density of less than 0.4.

The *z* axis of Eq. 1 parallels the Fe-O bond within 5°. The unpaired electron in the d_{yz} orbital produces a large negative *x* component of the spin-dipolar part of the *A* tensor, $A_{SD}/g_n\beta_n$ = (−24.4, +8.7, +15.7) T, and together with the isotropic contact term gives the larger *A* value along *x*. Owing to the large energy gaps between the narrowly spaced (d_{xz}, d_{yz}) pair and the other 3d levels, spin-orbit coupling acts primarily between d_{xz} and d_{yz} and produces (i) an effective g value g_x = 2(1 − ζ/Δ) (g_x and g_y are not affected) and (ii) an orbital contribution to the magnetic hyperfine interaction, $A_{L,x}I_xS_x$, where $A_{L,x}$ = $P(g_x - 2)$ ≈ −13 T. Substitution of ζ = 260 cm^{-1} and Δ_{DFT} = 2000 cm^{-1} gives the observed value g_x = 1.74 and $A_{L,x}$ = −13 T (*19*).

Complex **3** thus appears to have a low-spin d^3 configuration in contrast to a recently described isoelectronic manganese(IV)-oxo complex (*20*), which is high spin, *S* = 3/2, by EPR criteria. Our particular interest in TAML-catalyzed reactions pertains to their proven potential in green chemistry. The present work provides an opportunity to elucidate elementary steps in the catalytic cycles of TAML activators by taking advantage of the high purity of **3**, with its various prominent spectroscopic signatures, as a stoichiometric reagent.

References and Notes

1. I. G. Denisov, T. M. Makris, S. G. Sligar, I. Schlichting, *Chem. Rev.* **105**, 2253 (2005).
2. E. I. Solomon *et al.*, *Chem. Rev.* **100**, 235 (2000).
3. M. Costas, M. P. Mehn, M. P. Jensen, L. Que Jr., *Chem. Rev.* **104**, 939 (2004).
4. D. N. Harischandra, R. Zhang, M. Newcomb, *J. Am. Chem. Soc.* **127**, 13776 (2005).
5. M. Newcomb *et al.*, *J. Am. Chem. Soc.* **128**, 4580 (2006).
6. N. Aliaga-Alcalde *et al.*, *Angew. Chem. Int. Ed.* **44**, 2908 (2005).
7. A. Chanda *et al.*, *J. Inorg. Biochem.* **100**, 606 (2006).
8. A. Ghosh *et al.*, *J. Am. Chem. Soc.* **127**, 2505 (2005).
9. D. Banerjee *et al.*, *Angew. Chem. Int. Ed.* **45**, 3974 (2006).
10. S. S. Gupta *et al.*, *Science* **296**, 326 (2002).
11. M. J. Bartos *et al.*, *Coord. Chem. Rev.* **174**, 361 (1998).
12. M. R. Bukowski *et al.*, *Science* **310**, 1000 (2005).
13. J.-U. Rohde *et al.*, *Science* **299**, 1037 (2003).
14. Materials and methods are available as supporting material on *Science* Online.
15. The 3d spin-flip transition $(d_{yz})^{\beta} \rightarrow (d_{yz})^{\alpha}$ is higher in energy than the ligand-to-metal transition $(\pi_p)^{\beta} \rightarrow (d_{yz})^{\beta}$, where π_p is the redox-active orbital of the phenyl moiety, owing to the large (d_{xz}, d_{yz}) − d_{xy} splitting.
16. J. T. Groves, *J. Inorg. Biochem.* **100**, 434 (2006).
17. A. Decker, M. D. Clay, E. I. Solomon, *J. Inorg. Biochem.* **100**, 697 (2006).
18. F. Neese, *J. Inorg. Biochem.* **100**, 716 (2006).
19. The free-ion value for the one-electron spin-orbit coupling constant ζ is 580 cm^{-1} (*21*). Given the high degree of covalency of **3**, the adopted ζ is not unreasonable. $P = 2g_n\beta\beta_n\langle r^{-3}\rangle_{3d}$ has been taken as 50 T.
20. T. H. Parsell, R. K. Behan, M. T. Green, M. P. Hendrich, A. S. Borovik, *J. Am. Chem. Soc.* **128**, 8728 (2006).
21. J. Bendix, M. Brorson, C. E. Schaffer, *Inorg. Chem.* **32**, 2838 (1993).

22. The work was supported by NIH grants EB001475 (E.M.) and GM-38767 (L.Q.), the Heinz Endowments (T.J.C.), Environmental Protection Agency (grant RD 83 to T.J.C.), and the Heinz Foundation for the Teresa Heinz Scholars for Environmental Research award (to D.B.). XAS data were collected on beamline X3B at the National Synchrotron Light Source, which is supported by the U.S. Department of Energy and the National Institutes of Health. We also thank M. Costas and A. Ghosh for valuable suggestions.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1133417/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
Tables S1 to S6
References

3 August 2006; accepted 12 December 2006
Published online 21 December 2006;
10.1126/science.1133417
Include this information when citing this paper.

Small Phytoplankton and Carbon Export from the Surface Ocean

Tammi L. Richardson[1]* and George A. Jackson[2]

Autotrophic picoplankton dominate primary production over large oceanic regions but are believed to contribute relatively little to carbon export from surface layers. Using analyses of data from the equatorial Pacific Ocean and Arabian Sea, we show that the relative direct and indirect contribution of picoplankton to export is proportional to their total net primary production, despite their small size. We suggest that all primary producers, not just the large cells, can contribute to export from the surface layer of the ocean at rates proportional to their production rates.

Carbon export from the oceanic surface layer is controlled by biological transformations that occur within the pelagic food web (*1*, *2*). Autotrophic picoplankton (<2 μm in diameter) often dominate primary production in these regions but are believed to contribute relatively little to carbon export from surface layers because of their small sizes, slow sinking rates, and rapid utilization in the microbial loop. Large, rapidly sinking phytoplankton, such as diatoms, are believed to control carbon flux from upper ocean layers, and their contributions to export are believed to be disproportionately larger than their contributions to total primary production (*1*). Here we ask whether the contributions of picoplankton to carbon fluxes from the surface ocean are, in fact, disproportionately low.

Using results from inverse and network analyses of U.S. Joint Global Ocean Flux Study (JGOFS) data from the equatorial Pacific (EqPac) and Arabian Sea, we show that the relative contributions of various phytoplankton size classes to carbon export are proportional to their contributions to total net primary production (NPP) (*3*) (Fig. 1). Potential export pathways include direct routes, such as aggregation and incorporation into settling detritus, and indirect export through the consumption of picoplankton aggregates by organisms at higher trophic levels. The network analysis used in these studies (*4*) estimates the total particulate organic carbon (POC) flux that originated from each input source, here the NPP of each phytoplankton size class. In the EqPac, for example, picoplankton contributed 70% or more of the total NPP measured during JGOFS EqPac time series cruises and were responsible for 87% of POC export via detritus and 76% of carbon exported through the mesozooplankton (*5*). Mesozooplankton consumed both picoplankton-derived detritus and picoplankton-fed micrograzers. Similarly, at a northern station in the Arabian Sea during the Northeast Monsoon, picoplankton NPP was 86% of the total and was the source for 97 and 75% of the total carbon exported from the euphotic zone via the POC and mesozooplankton pathways, respectively (*6*). The relative contributions of the picoplankton to carbon export decreased where they produced relatively less of the total NPP but were nonetheless substantial (Fig. 1 and Table 1).

We believe that the contributions of picoplankton to carbon flux from oceanic surface layers have been overlooked because of assumptions that are made about cell size, sinking dynamics, and the trophic pathways of picoplankton through ecosystems. The seminal paper

[1]Marine Sciences Program and Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, USA. [2]Department of Oceanography, Texas A&M University, MS 3146, College Station, TX 77843–3146, USA.

*To whom correspondence should be addressed. E-mail: richardson@biol.sc.edu

by Smayda (7) documented the tendency of larger cells to sink faster. However, there are at least three potential mechanisms for increasing the effective size of picoplankton and their removal rates from the euphotic zone. First, the aggregation of picoplankton cells into larger detrital particles (8, 9) enhances their vertical settling velocities and resulting export fluxes (10–12). A simple model of aggregation and cell sinking for phytoplankton of varying diameters (3) (Fig. 2) shows that, at peak velocities, the average settling rate for aggregated cells from 1 to 30 μm in diameter is similar, although aggregation does not substantially increase the settling rates of the largest cells modeled (100 μm). We conclude that an individual alga does not need to be large to sediment out, even though settling rates are generally faster for single cells that are larger. Aggregation may be enhanced if cells are nutrient-depleted (13), and settling may be enhanced by the incorporation of mineral matter (14–16).

The inclusion of small phytoplankton in marine snow aggregates is well documented (2, 17–19). Aggregates of intact *Emiliania huxleyi* were first found in sediment traps deployed for short time periods (24 to 48 hours) in the North Sea (20). In the northeast Atlantic, water samples captured with bottles showed concentrations of *Synechococcus*-like cyanobacteria on aggregates that were 3000 to 17,000 times higher in concentration than that seen in free-living forms (18). Picoplankton aggregates have also been observed in sediment traps equipped with acrylamide gels during short-term (48 hours) deployments in high-nutrient, low-chlorophyll waters off New Zealand (19). It is noteworthy that observations of aggregates of small phytoplankton in sediment traps have been in traps deployed for short time periods, which will minimize loss of material and maximize potential identification of constituent organisms as compared to longer-term deployments (21).

Although we did not explicitly include the process of aggregation in our EqPac and Arabian Sea analyses, we did allow ungrazed picoplankton production to flow directly to the detritus pool, where it could settle out as detritus or be consumed by larger zooplankton. Models of food webs and biogeochemical cycling (22–25) generally assume that all picoplankton production is cycled through the microbial loop and that none sinks from the euphotic zone directly. This follows the classic analysis of pelagic community structure and nitrogen fluxes by Michaels and Silver (1), who assumed that "only large particles, with consequently high potential sinking rates, can exit the euphotic zone" and that "large colonies of smaller algae" decompose (and are presumably recycled) within the euphotic zone. Implicit in these food web structures is the assumption that growth and grazing in picoplankton-dominated open oceans are in balance over long (annual) time scales (26). However, our studies (5, 6) and those of others (27) have shown that microzooplankton grazing does not always balance picoplankton growth on shorter time scales (weeks to months). Food web models that force all picoplankton production through the microbial loop and do not allow direct picoplankton export may, therefore, be misleading. The conclusion of Michaels and Silver (1) that "picoplankton, the dominant producers in the model, contribute little to the sinking material" was based on [illegible] that was

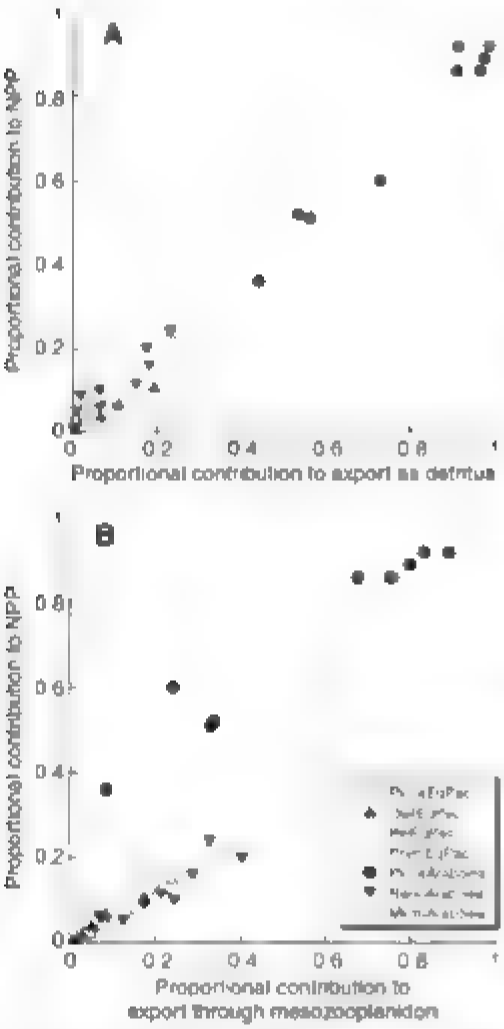


Fig. 1. Proportional contributions of varying phytoplankton groups or size classes (picoplankton, diatoms, pelagophytes, and prymnesiophytes for the EqPac study and pico-, nano-, and microphytoplankton for the Arabian Sea) to NPP versus their proportional contributions to export as detritus (**A**) or through consumption of mesozooplankton (**B**). Proportional contributions were calculated as NPP or export due to the size class/total NPP or export flux.

Table 1. Direct and indirect contributions of picoplankton to carbon export from the euphotic zone for two cruises in the EqPac and at three stations (N7, S2, and S11) during three seasons in the Arabian Sea. Abbreviations are as follows: TS1, time series 1 (March to April 1992); TS2 Early, time series 2 cruise (early October 1992); TS2 Mid, midway through TS2 (mid-October 1992); TS2 Late, late October 1992; NEM, Northeast Monsoon; SIM, Spring Intermonsoon; and SWM, Southwest Monsoon.

Model	Export pathway	Carbon export due to picoplankton (mmol of C m^{-2} day^{-1})	% Contribution of picoplankton to total carbon export flux
	EqPac		
TS1	POC	1.7	87
	Mesozooplankton	10.4	76
TS2 Early	POC	0.7	57
	Mesozooplankton	11.5	60
TS2 Mid	POC	1.5	43
	Mesozooplankton	11.0	50
TS2 Late	POC	1.7	63
	Mesozooplankton	11.2	63
	Arabian Sea		
NEM N7	POC	1.0	97
	Mesozooplankton	11.5	75
NEM S2	POC	3.8	56
	Mesozooplankton	5.0	33
NEM S11	POC	4.3	99
	Mesozooplankton	11.1	83
SIM N7	POC	3.7	73
	Mesozooplankton	3.4	24
SIM S2	POC	5.9	92
	Mesozooplankton	2.4	89
SIM S11	POC	2.8	98
	Mesozooplankton	11.0	80
SWM N7	POC	2.3	91
	Mesozooplankton	1.7	68
SWM S2	POC	13.8	53
	Mesozooplankton	1.4	34
SWM S11	POC	10.5	44
	Mesozooplankton	0.8	1

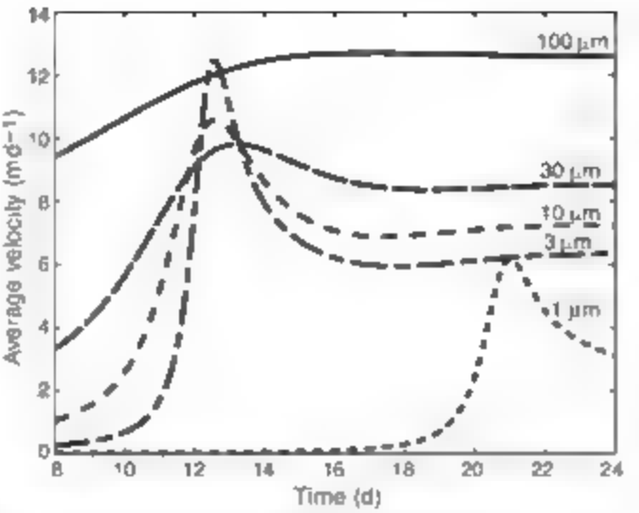


Fig. 2. Average sinking velocity versus time (in days) for phytoplankton cells with diameters of 1, 3, 10, 30, and 100 μm. A cell density was assigned that is consistent with the velocity versus cell diameter graph of Smayda (7) for 3-, 10-, 30-, and 100-μm cells. Because 1-μm cells are outside the range analyzed by Smayda, a density difference of 0.0725 g cm^{-3}, characteristic of *Synechococcus lividus*, was used (3). Cells grow exponentially in a mixed layer of constant thickness, and there was no light or nutrient limitation. Algal cells collide to form aggregates at rates that depend on their abundances as well as sizes. Large cells settle faster than small ones. With time, cell concentrations increase, causing an increase in the fraction of material in aggregates. The resulting increase in average particle size leads to an increase in average settling speed. Peaks in total cell concentration occur when the enhanced losses due to settling balance the gains due to cell division. The maximum average settling rate for particles formed from the 1-μm cells is not substantially different from that for particles formed from the 30-μm cells.

structured in such a way that no other conclusion could be reached.

A second pathway for the accelerated sinking of picoplankton-derived material is through mesozooplankton grazing. Picoplankton are usually considered to be too small to be captured effectively by larger grazers such as copepods (*28*), but the increase in effective size through aggregation makes them available and thus can enhance their export from surface layers through their incorporation into fast-sinking fecal pellets (*9*). When mesozooplankton consume picoplankton-containing aggregates, the picoplankton carbon short-circuits the microbial loop and results in higher than expected efficiency of carbon transfer from the euphotic zone (*18*, *29*, *30*). For animals such as salps that are capable of feeding on particles as small as 1 μm, aggregation is unnecessary to produce fecal pellets. The aggregates described in (*19*) were thought to originate in salp or other tunicate fecal material. The incorporation of picoplankton-sized particles into the rapidly sinking mucous nets and feces of filter feeders such as salps or pteropods provides an additional route for picoplankton removal (*2*, *31*).

We propose, therefore, that the conventional view of carbon cycling, in which picoplankton carbon is recycled within the microbial loop and only larger phytoplankton is exported, should be revised to include the alternative pathways for picoplankton carbon cycling discussed above. This alternative view relies on aggregation as a mechanism for both direct sinking (the export of picoplankton as POC) and mesozooplankton- or large filter feeder–mediated sinking of picoplankton-based production, but we believe that the evidence for both these processes under an array of environmental conditions is well established. Despite their small size, picoplankton may contribute more to oceanic carbon export than is currently recognized, and these contributions should be considered in current models of trophic dynamics in the ocean.

References and Notes

1. A. F. Michaels, M. W. Silver, *Deep-Sea Res.* **35**, 473 (1988).
2. L. Legendre, J. Le Fèvre, *Aquat. Microb. Ecol.* **9**, 69 (1995).
3. Materials and methods are available as supporting material on *Science Online*.
4. R. E. Ulanowicz, *NETWRK4: A Package of Computer Algorithms to Analyze Ecological Flow Networks* (Chesapeake Biological Laboratory, Univ. of Maryland, Solomons, MD, 1987).
5. T. L. Richardson, G. A. Jackson, H. W. Ducklow, M. R. Roman, *Deep Sea Res. I* **51**, 1245 (2004).
6. T. L. Richardson, G. A. Jackson, H. W. Ducklow, M. R. Roman, *Deep-Sea Res. II* **53**, 555 (2006).
7. T. J. Smayda, *Oceanogr. Mar. Biol. Ann. Rev.* **8**, 353 (1970).
8. P. Albertano, D. D'Oronma, E. Caputi, *J. Plankton Res.* **19**, 1405 (1997).
9. K. Olli, A.-S. Heiskanen, *J. Mar. Syst.* **23**, 165 (1999).
10. G. A. Jackson, *Deep-Sea Res. I* **37**, 1197 (1990).
11. G. A. Jackson, *Deep-Sea Res. I* **48**, 95 (2001).
12. G. A. Jackson, A. M. Waite, P. W. Boyd, *Geophys. Res. Lett.* **32**, L13067 (2005).
13. M. E. Klut, J. G. Stockner, *Can. J. Fish. Aquat. Sci.* **48**, 1092 (1991).
14. H. D. R. Gomes, J. I. Goes, A. H. Pasulekar, *J. Plankton Res.* **14**, 1307 (1992).
15. C. Klaas, D. E. Archer, *Global Biogeochem. Cycles* **16**, 1116 (2002).
16. R. A. Armstrong, C. Lee, J. I. Hedges, S. Honjo, S. G. Wakeham, *Deep-Sea Res. II* **49**, 219 (2002).
17. M. Revelante, M. Gilmartin, *J. Exp. Mar. Biol. Ecol.* **146**, 217 (1991).
18. R. S. Lampitt, K. F. Wishner, C. M. Turley, M. V. Angel, *Mar. Biol.* **116**, 689 (1993).
19. A. M. Waite, K. A. Safi, A. Hall, S. D. Nodder, *Limnol. Oceanogr.* **45**, 87 (2000).
20. G. C. Cadée, *Mar. Ecol. Prog. Ser.* **24**, 193 (1985).
21. K. O. Buesseler *et al.*, *Report of SCOR Working Group 116* (International Council of Scientific Unions, Woods Hole, MA, 2006).
22. R. B. Rivkin *et al.*, *Science* **272**, 1163 (1996).
23. P. W. Boyd, P. P. Newton, *Deep-Sea Res. I* **46**, 63 (1999).
24. E. A. Laws, *Global Biogeochem. Cycles* **14**, 1231 (2000).
25. F. Chai, R. C. Dugdale, T.-H. Peng, F. P. Wilkerson, R. T. Barber, *Deep-Sea Res. II* **49**, 2713 (2002).
26. M. R. Landry *et al.*, *Limnol. Oceanogr.* **42**, 405 (1997).
27. M. R. Roman *et al.*, *Deep-Sea Res. II* **49**, 175 (2002).
28. P. Nival, S. Nival, *Limnol. Oceanogr.* **21**, 24 (1976).
29. U. Passow, A. L. Alldredge, *J. Plankton Res.* **21**, 2203 (1999).
30. R. Mari, F. Rassoulzadegan, *Mar. Ecol. Prog. Ser.* **279**, 13 (2004).
31. T. T. Noji *et al.*, *J. Plankton Res.* **19**, 863 (1997).

32. This work was supported by NSF (grants OCE 0097296 and OCE 0352127 to G.A.J.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5813/838/DC1
Materials and Methods
References

4 August 2006; accepted 20 December 2006
10.1126/science.1133471

Integration of TGF-β and Ras/MAPK Signaling Through p53 Phosphorylation

Michelangelo Cordenonsi, Marco Montagner, Maddalena Adorno, Luca Zacchigna, Graziano Martello, Anant Mamidi, Sandra Solngo, Sirio Dupont, Stefano Piccolo*

During development and tissue homeostasis, cells must integrate different signals. We investigated how cell behavior is controlled by the combined activity of transforming growth factor–β (TGF-β) and receptor tyrosine kinase (RTK) signaling, whose integration mechanism is unknown. We find that RTK/Ras/MAPK (mitogen-activated protein kinase) activity induces p53 N-terminal phosphorylation, enabling the interaction of p53 with the TGF-β–activated Smads. This mechanism confines mesoderm specification in *Xenopus* embryos and promotes TGF-β cytostasis in human cells. These data indicate a mechanism to allow extracellular cues to specify the TGF-β gene expression program.

Cross-talk between signaling pathways is necessary to effect efficient and fine-tuned regulatory control over metazoan development and physiology. TGF-β and receptor tyrosine kinase (RTK) ligands are pleiotropic cytokines affecting several aspects of cell behavior, ranging from differentiation and proliferation to movement and survival (*1*, *2*). Previous work has shown that these signaling pathways are integrated: The Ras/MAPK cascade, which is downstream of RTK signaling, affects TGF-β–induced mesoderm development in vertebrate

embryos and growth arrest in mammalian adult cells (*1*, *3*, *6*). However, the mechanisms underlying this partnership have not been elucidated.

Smad2/3 and the tumor suppressor protein p53 physically interact and jointly regulate the transcription of several TGF-β target genes (*7*). p53 is activated by multiple stimuli through posttranslational modifications (*8*). Hence, p53 activation might serve to convey cues from extracellular signals within the TGF-β gene-expression program.

To investigate whether p53 acts as an integration node between Ras-MAPK and TGF-β pathways, we carried out loss-of-function studies in *Xenopus* embryos, where pluripotent cells of the animal pole (animal cap) can differentiate into mesoderm by the combined action of these signals. Endogenous p53 was depleted by microinjecting p53 antisense morpholino oligonucleotides (p53-MO) (fig. S1A). Control or p53-depleted animal cap cells were treated with combinations of fibroblast growth factor (FGF) and Activin proteins. FGF enhanced Activin-mediated induction of mesoderm markers (Fig. 1A, lanes 3 and 4), but this cooperation was lost in p53-depleted cells (Fig. 1A, lanes 7 and 8). Moreover, both p53 and FGF are required for the induction of a panel of Activin target genes (fig. S1B). These results are consistent with p53 being required downstream of FGF to foster TGF-β gene responses.

Upon overexpression of p53 in *Xenopus* animal cap cells, p53 cooperates with endogenous Smads to induce mesoderm markers (*7*) (fig. S1C); this p53 activity is counteracted by treatment of animal caps with the FGF-receptor inhibitory compound SU 5402 or expression of a dominant-negative version of Raf, a critical component of the Ras-MAPK pathway (fig. S1, D and E). This suggests that endogenous RTK signaling promotes the mesoderm-inducing activities of p53. Moreover, p53 and FGF cooperate in mesoderm induction, as assayed by the induction of ectopic tail structures in whole embryos (fig. S1, F to J). To address the biochemical basis of this link, we treated human HepG2 cells with combinations of FGF and Activin proteins and then purified p53 from corresponding nuclear lysates by DNA-affinity purification (DNAP). As shown in Fig. 1B, treatment with FGF efficiently promotes the association of p53 and TGF-β activated Smad2 within the same complex.

To gain insight into this process, we defined the structural determinants of p53 that are relevant for Smad binding. p53 binds recombinant Smad3 through the p53 N-terminal domain (fig. S2A), this segment carries several Ser/Thr residues (fig. S2B), whose phosphorylation has been implicated in p53 activation upon DNA damage (*8*). We found that endogenous FGF signaling also promotes phosphorylation of these residues in *Xenopus* embryos (fig. S2C).

To address whether N-terminal phosphorylation plays a causative role in guiding p53 activity toward the TGF-β pathway, we compared wild-type and N-mut Xp53, in which the N-terminal Ser/Thr residues have been replaced by Ala, for their ability to rescue TGF-β gene responses in p53-depleted *Xenopus* embryos. Mutation of the N-terminal phosphorylation sites severely impairs p53 mesoderm-inducing ability (Fig. 1C). Similar results were obtained with wild-type and N-mut mammalian p53 (fig. S2D). At the biochemical level, only wild-type and not N-mut p53 can complex with Smad2 (Fig. 1D and fig. S2E), indicating that this interaction, rather than being constitutive, must be enabled by p53 N-terminal phosphorylation. This appears to be a peculiar requirement, because wild-type and N-mut p53 display similar stability, transactivation capacity, and apoptosis-inducing activity (fig. S2, F to H). These results indicate that phosphorylation of N-terminal Ser/Thr residues is relevant for coupling p53 activity to Smad responses.

To investigate the relevance of p53 N-terminal phosphorylation for the activation of the TGF-β cytostatic program in human cells, we established a p53-complementation assay using the p53-null H1299 human lung cancer cell line (*9*). These cells retain an intact TGF-β transduction cascade and yet are unable to activate the TGF-β cytostatic program (fig. S3). Robust TGF-β mediated induction of the cyclin-dependent kinase (CDK) inhibitors $p21^{Waf1}$ and $p15^{Ink4b}$ is rescued by adding back wild-type p53, whereas expression of p53-N-mut fails to do so (fig. S3B).

To identify the residues that must be phosphorylated in vivo to enable p53 Smad cooperation, we refined our analysis by comparing wild-type p53 with p53 mutants carrying Ala

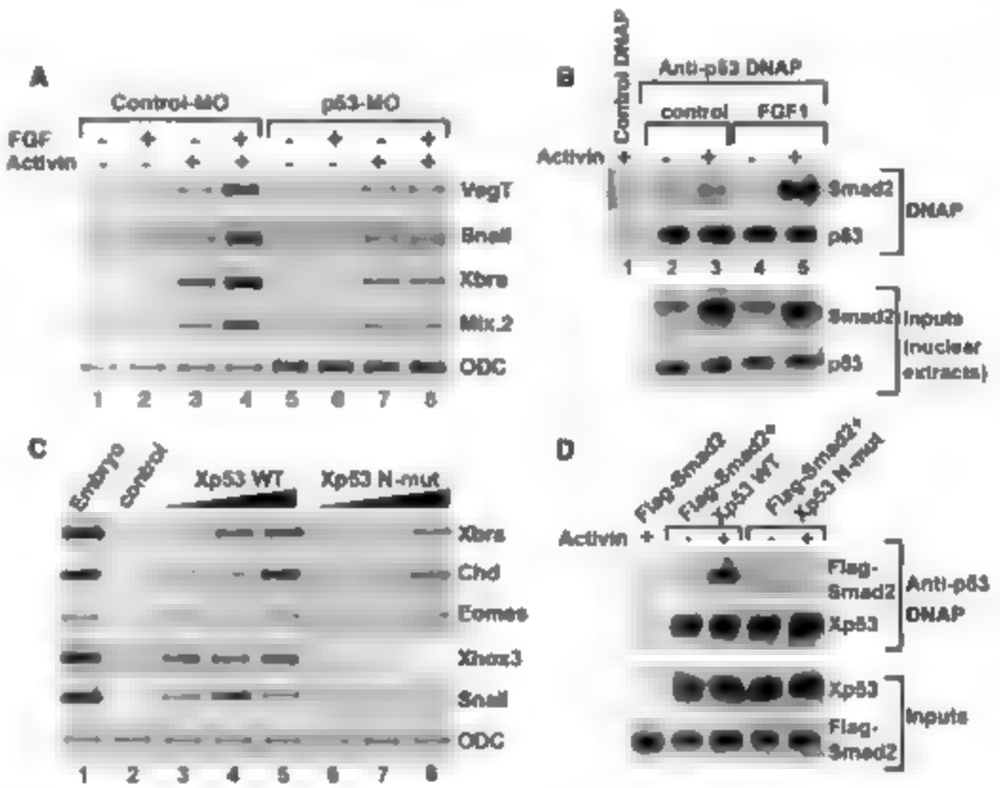


Fig. 1. FGF potentiates TGF-β gene responses in a p53-dependent manner. (**A**) Reverse transcription polymerase chain reaction (RT-PCR) analysis for mesodermal markers (*VegT*, *Snail*, *Xbra*, and *Mix.2*) of animal caps explanted from *Xenopus* embryos injected with control-MO or p53-MO (40 ng). Where indicated, explants were treated with FGF1 (25 ng/ml) and Activin (6 ng/ml) and cultured for 2 hours before harvesting. The samples injected with p53-MO (lanes 5 to 8) were subjected to two additional PCR cycles for all the markers analyzed, in order to visualize residual mesodermal gene activations triggered by Activin in the absence of p53. ODC (ornithine decarboxylase) serves as a loading control. (**B**) FGF and Activin cooperate for the formation of a Smad2/p53 protein complex. Western blot analyses for endogenous p53 and Smad2 of protein complexes purified by DNA affinity purification (DNAP) with a wild-type (anti-p53 DNAP, lanes 2 to 5) or mutant (control DNAP, lane 1) biotinylated p53-consensus probe (*7*). (**C**) RT-PCR analyses for mesodermal marker genes induced in animal cap explants by *wild-type* (wt) or *N-mut Xp53* mRNAs. *Xenopus* embryos were injected with p53-MO (40 ng) and different doses (10 pg, 30 pg, and 90 pg) of mRNAs encoding for the Xp53 isoforms. (**D**) Interaction of Xp53 and Smad2 requires p53 N-terminal phosphorylation. Extracts from *Xenopus* embryos injected with combinations of mRNAs (100 pg each) for Flag-tagged Smad2, wild-type or N-mut *Xenopus* p53, and Activin were precipitated by anti-p53 DNAP. The panels show Western blots for Smad2 and Xp53.

Department of Medical Biotechnologies, Section of Histology and Embryology, University of Padua, Padua, Italy.

*To whom correspondence should be addressed. E-mail: piccolo@cu.bio.unipd.it

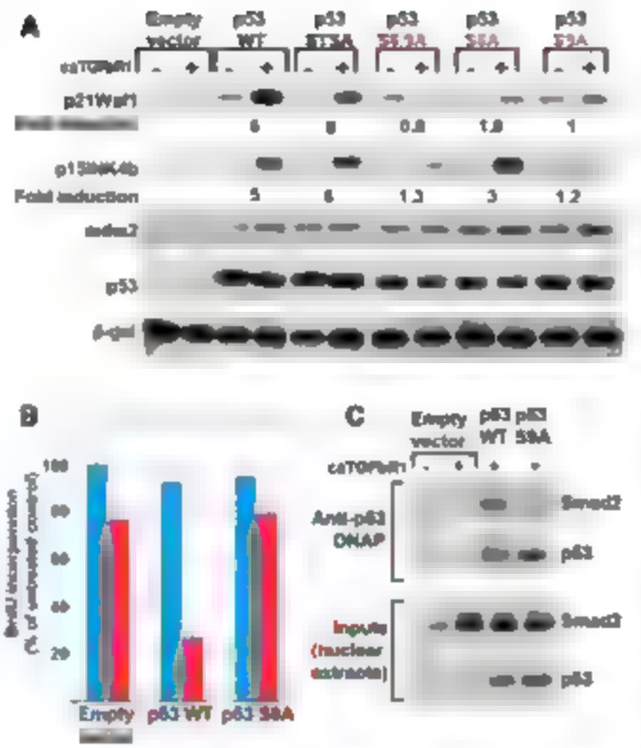


Fig. 2. Requirement of p53 Ser[9] phosphorylation for activation of the TGF-β cytostatic program. (**A**) Phosphomutant-p53 isoforms were tested for the ability to rescue TGF-β responsiveness in H1299 cells in comparison with wild-type mouse p53. p53 S15A carries Ala substitutions in Ser[15], Thr[18] and Ser[20]; p53S6,9A, S6A, or S9A carry Ala substitutions in Ser[6] and/or Ser[9]. Transfection and analysis of H1299 were as in fig. S3B. Fold inductions are the ratio of $p21^{Waf1}$ or $p15^{Ink4b}$ expression in the presence or absence of TGF-β stimulation, normalized on p53 levels. (**B**) Wild-type, but not p53S9A, rescues TGF-β–induced growth arrest in H1299 cells. Cells were transfected with the indicated p53 expression constructs as in (A) and assayed for BrdU incorporation. Columns represent the number of BrdU positive cells in the absence (cyan) or presence (red) of TGF-β stimulation, relative to the number of proliferating cells in the unstimulated control. (**C**) The interaction of p53 and Smad2 requires Ser9 phosphorylation. Nuclear extracts from H1299 cells transfected either with wild-type mouse p53 or p53S9A were precipitated by anti-p53 DNAP. The panels show Western blots for Smad2 and p53.

substitutions in (i) Ser[15], Thr[18], and Ser[20], (ii) Ser[6] and Ser[9], or (iii) individual residues. As shown in Fig. 2A, all of these p53 isoforms similarly rescued the expression of mdm2, a TGF-β independent p53 target, as well as the basal levels of $p21^{Waf1}$. Moreover, Ser[15], Thr[18] and Ser[20] were not required for inducibility of $p21^{Waf1}$ and $p15^{Ink4b}$ by TGF-β signaling. In contrast, phosphorylation of Ser[9] and Ser[6] was relevant for Smad cooperation. In line with these results, Ser[9] phosphorylation was required to fully empower the mesoderm-inducing properties of p53 in *Xenopus* embryos (fig. S4). To extend these observations to TGF-β induced cytostasis, we measured incorporation of bromodeoxyuridine (BrdU) in parental (p53-null) and p53-reconstituted H1299 cells. As shown in Fig. 2B, only wild-type p53, but not p53S9A, could rescue TGF-β dependent growth arrest. Mechanistically, this is due to an impaired ability of p53S9A to complex with Smad2 (Fig. 2C). We then investigated whether endogenous Ras/MAPK signaling is relevant for p53 phosphorylation in Ser[6]/Ser[9] (P-Ser). H1299 cells carry an activating mutation in N-Ras, leading to constitutive MAPK signaling (9). We found that inhibition of MAPK kinase (MEK), an effector of Ras upstream of MAPK, causes specific inhibition of P-Ser[6] and P-Ser[9] levels, with concomitant loss of TGF-β mediated $p21^{Waf1}$ induction in p53-reconstituted H1299 cells (fig.

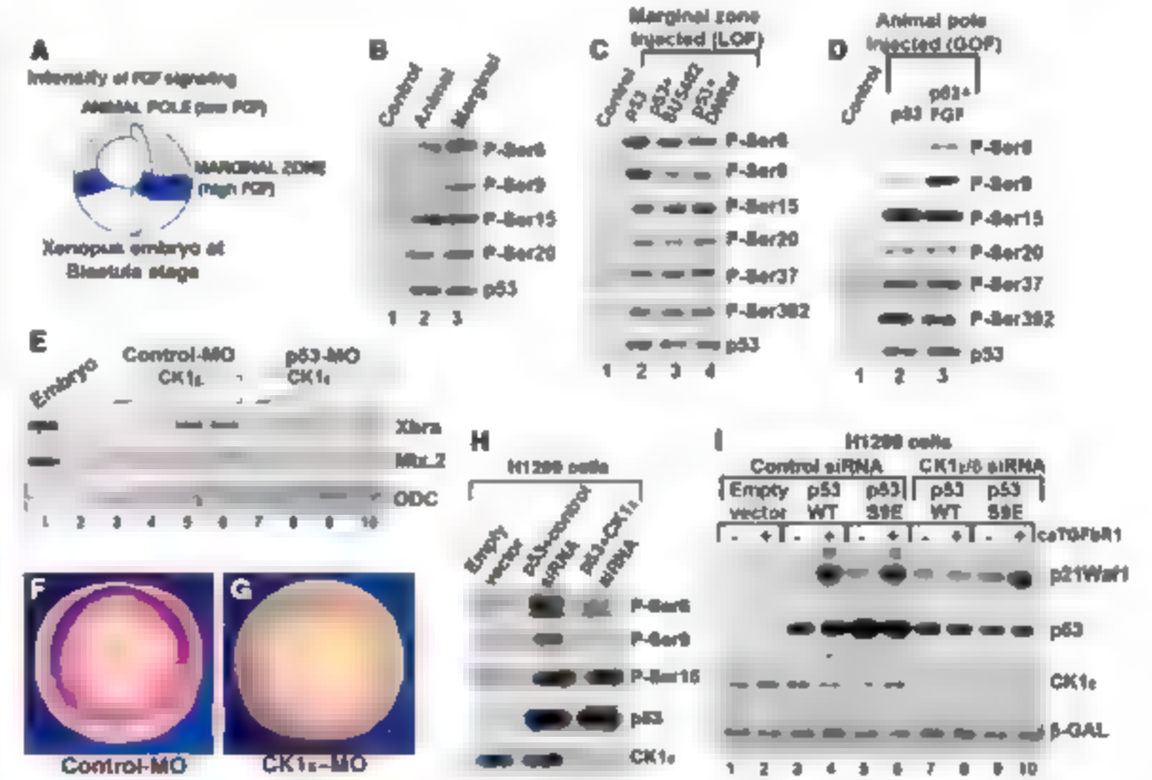


Fig. 3. FGF phosphorylates p53 on Ser[6] and Ser[9] through CK1ε/δ. (**A**) Schematic diagram showing the distribution of FGF/MAPK signaling in the *Xenopus* embryo at late blastula stage (*10*). (**B** to **D**) Analysis of the phosphorylation status of human p53 (100 pg) injected in *Xenopus* embryos. p53 was purified by immunoprecipitation, and phosphoresidues were detected by Western blot. (**B**) *p53* mRNA was injected in the animal pole or in the marginal zone of *Xenopus* embryos. p53 phosphorylation on Ser[6] and Ser[9] is enriched in the marginal zone, where FGF signaling is stronger. (**C**) *p53* mRNA was injected in the marginal zone region alone or in combination with *DN-Raf* mRNA (1 ng). When indicated, injected embryos were cultivated in the presence of the FGFR1 inhibitor SU5402 (60 μM). (**D**) FGF enhances p53 phosphorylation of Ser[9] and, to a minor extent, of Ser[6]. *p53* mRNA was injected in the animal pole region alone, or in combination with *eFGF* mRNA (0.8 pg). Consistent results were obtained on endogenous p53 in human cells (fig. S5B). (**E**) CK1ε induces expression of mesodermal genes in a p53-dependent manner. *Xenopus CK1ε* mRNA was injected at four different doses (200 pg, 400 pg, 800 pg, and 1.6 ng) together with control-MO or p53-MO (40 ng). Animal cap explants were dissected at late blastula stage as in Fig. 1A. Additional markers are shown in fig. S8A. (**F** and **G**) Whole-mount in situ hybridizations for the pan-mesodermal marker *Xbra* in control-MO and CK1ε-MO–injected embryos. (**H**) CK1ε/δ are required for Ser[6] and Ser[9] phosphorylation in human cells. The panel shows Western blot analysis of p53 phosphorylation in H1299 cells. Wild-type mouse p53 was transfected in combination with control-siRNA (small interfering RNA) or anti-CK1ε/δ siRNA. CK1ε depletion was monitored by Western blot. (**I**) CK1ε/δ is required for the TGF-β cytostatic program through p53 Ser[9] phosphorylation.

S5A). Hence, p53 phosphorylation in Ser[9] and Ser[6] serves as integration node in the cross-talk between Ras MAPK and TGF-β.

This prompted us to consider the possibility that although p53 is a ubiquitous protein, FGF might spatially pattern p53's activity. In *Xenopus*, expression of different FGFs (eFGF, FGF3, and FGF8) is enriched in the marginal zone of the embryo, from which the mesoderm emerges, whereas lower FGF activity is present in the animal pole (*10*) (Fig. 3A). Using phosphospecific antibodies, we found that kinase activities targeting Ser[9] and Ser[6] are localized in the marginal zone; in contrast, phosphorylation on other residues appears constitutive (Fig. 3B). To determine whether endogenous FGF signaling is responsible for this graded p53 phosphorylation along the animal-vegetal axis, embryos were treated with the FGF-receptor inhibitor SU 5402 or injected with *DN Raf* mRNA. Blockade of FGF signaling causes specific down-regulation of P-Ser[9] and P-Ser[6] (Fig. 3C). Conversely, ectopic FGF expression in animal cap cells specifically raises P-Ser[6] and P-Ser[9] levels (Fig. 3D). Similarly, at the biochemical level, FGF is required for p53-Smad2 interaction because the formation of this complex is inhibited by SU 5402 (fig. S6). However, introduction of Ser to Glu phosphomimicking substitutions in Ser[6] and Ser[9] (p53S6,9E), renders p53 able to complex with Smad2 in an FGF-independent manner (fig. S6). Together, the results indicate that FGF patterns the phosphorylation status of p53 in the embryo, restricting its cooperation with TGF-β to the prospective mesoderm.

Next, we wished to gain insight into the kinase responsible for inducing p53 phosphorylation in response to FGF-Ras-MAPK signaling. Both Ser[6] and Ser[9] conform to a CK1 consensus. There are seven mammalian CK1 genes, but p53 has been shown to associate specifically with CK1ε and CK1δ (*11*). In *Xenopus* embryos, inhibition of these kinases with dominant-negative (DN-) CK1δ (DN-CK1δ) (*12, 13*) antagonizes FGF-mediated Ser[6] and Ser[9] phosphorylation (fig. S7). Biologically, increasing levels of CK1ε promote mesoderm induction in a p53-dependent manner (Fig. 3E and fig. S8); conversely, loss-of-CK1ε by microinjection of DN-CK1ε or CK1ε morpholino inhibits endogenous and p53-mediated mesodermal gene expression (Fig. 3, F and G, and fig. S9). Thus, CK1ε lies downstream of FGF to promote p53 phosphorylation and Smad cooperation in *Xenopus* mesoderm development.

We next investigated the relevance of CK1ε δ-mediated p53 phosphorylation on the activation of the TGF-β cytostatic program in human cells. To this end, p53-reconstituted H1299 cells were transfected with siRNAs to deplete endogenous CK1ε and CK1δ. CK1ε δ knockdown leads to down-regulation of P-Ser[9] and P-Ser[9] levels (Fig. 3H) and to loss of TGF-β–mediated $p21^{Waf1}$ induction (Fig. 3I, compare lanes 3 and 4 with lanes 7 and 8). By contrast, a phosphomimicking substitution of Ser[9] with Glu (p53S9E) renders p53 able to sustain TGF-β–mediated $p21^{Waf1}$ induction even in the absence of CK1ε δ (Fig. 3I, compare lane 4 with lane 8 and lane 6 with lane 10). Hence, p53S9E acts epistatically to CK1ε δ. This indicates the key role of p53 N-terminal phosphorylation as mediator of the positive effect of CK1ε δ in supporting TGF-β cytostatic responses.

We have established a role for p53 as signaling integrator, outside of its widely investigated response to genotoxic stress (*1*). We provide evidence that p53 activity, rather than stability, can be qualitatively patterned by RTK Ras-induced phosphorylation through CK1ε δ. This phosphorylation step enables a robust biochemical interaction of p53 with TGF-β activated Smads, leading to mesoderm induction in embryos and in human cells, to the deployment of the TGF-β cytostatic program.

These data establish a mechanistic link between three key regulators of cell proliferation that are dysregulated in human cancers: Ras, p53, and TGF-β. This could provide an explanation for the p53-dependent tumor-suppressive function of Ras MAPK reported in primary cells (*14, 15*). Activated Ras may well have general growth-promoting effects but, in the presence of wild-type p53, this would be balanced by the positive role played on p53 Smad cooperation that would sustain TGF-β growth control and thus limit neoplastic transformation.

References and Notes

1. L. Attisano, J. L. Wrana, *Science* **296**, 1646 (2002).
2. J. Schlessinger, *Cell* **103**, 211 (2000).
3. C. LaBonne, M. Whitman, *Development* **120**, 463 (1994).
4. R. A. Cornell, D. Kimelman, *Development* **120**, 453 (1994).
5. P. P. Hu *et al.*, *J. Biol. Chem.* **274**, 33381 (1999).
6. R. T. Bottcher, C. Niehrs, *Endocr. Rev.* **26**, 63 (2005).
7. M. Cardenoso *et al.*, *Cell* **113**, 301 (2003).
8. A. M. Bode, Z. Dong, *Nat. Rev. Cancer* **4**, 793 (2004).
9. T. Matsuzoni *et al.*, *Oncogene* **7**, 171 (1992).
10. C. LaBonne, B. Burke, M. Whitman, *Development* **121**, 1475 (1995).
11. U. Knippschild *et al.*, *Oncogene* **15**, 1727 (1997).
12. X. Zeng *et al.*, *Nature* **438**, 873 (2005).
13. G. Davidson *et al.*, *Nature* **438**, 867 (2005).
14. S. W. Lowe, E. Cepero, G. Evan, *Nature* **432**, 307 (2004).
15. E. Roper, W. Weinberg, F. M. Watt, H. Land, *EMBO Rep.* **2**, 145 (2001).

16. We thank G. Bressan and D. Volpin for discussion. We also thank G. Del Sal, M. Mechali, D. Lane, C. Niehrs, J. Graff, D. Morrison, and K. Vousden for gifts of plasmids or antibodies, G. Blandino for the gift of p53-reconstituted cells, and O. Wessely for comments. This work is supported by grants to S.P. from Associazione Italiana per la Ricerca sul Cancro, TELETHON-Italy GGP04030, MIUR (CoFin, FIRB), Agenzia Spaziale Italiana, Istituto Superiore di Sanità, and Swissbridge. A.M. is recipient of a European Union Marie Curie Research Training Network fellowship (Epiplard Carcinoma). M.C. was supported by a FIRC fellowship.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1135961/DC1
Materials and Methods
Figs. S1 to S9
References

5 October 2006; accepted 22 December 2006
Published online 18 January 2007
10.1126/science.1135961
Include this information when citing this paper.

Structure of the Prefusion Form of the Vesicular Stomatitis Virus Glycoprotein G

Stephane Roche, Felix A. Rey,* Yves Gaudin,† Stephane Bressanelli

Glycoprotein G of the vesicular stomatitis virus triggers membrane fusion via a low pH–induced structural rearrangement. Despite the equilibrium between the pre- and postfusion states, the structure of the prefusion form, determined to 3.0 angstrom resolution, shows that the fusogenic transition entails an extensive structural reorganization of G. Comparison with the structure of the postfusion form suggests a pathway for the conformational change. In the prefusion form, G has the shape of a tripod with the fusion loops exposed, which point toward the viral membrane, and with the antigenic sites located at the distal end of the molecule. A large number of G glycoproteins, perhaps organized as in the crystals, act cooperatively to induce membrane merging.

The Rhabdoviridae are enveloped bullet-shaped viruses that are widespread among a great variety of organisms, including plants, insects, fishes, mammals, reptiles, and crustaceans (*1*). This family includes vesicular stomatitis virus (VSV) as well as notable human pathogens, such as rabies virus (RV) and Chandipura virus (*2*).

The rhabdoviruses enter the cell via the endocytic pathway and subsequently fuse with a cellular membrane within the acidic environment of the endosome (*3*). Both receptor recognition and membrane fusion are mediated by a single transmembrane (TM) viral glycoprotein (G) that is trimeric and forms the spikes that protrude from the viral surface. The large ecto-

CNRS, Unité Mixte de Recherche (UMR) 2472, Institut National de la Recherche Agronomique (INRA), UMR 1157, Institut Fédératif de Recherche 115, Laboratoire de Virologie Moléculaire et Structurale, 91198, Gif sur Yvette, France.

*Present address: Département de Virologie, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris cedex 15, France.
†To whom correspondence should be addressed. E-mail: gaudin@vms.cnrs-gif.fr

domain of G (446 out of 495 amino acids for the VSV Indiana strain) is also the target of neutralizing antibodies, and the antigenic sites of G in both VSV and RV have been described in detail (*4*–*6*).

Similar to other viral fusion proteins, G undergoes a fusogenic structural transition during cell entry (*7*, *8*). As for influenza virus hemagglutinin (HA), flavivirus E protein, and Semliki Forest virus E1 protein, the conformational change is triggered at low pH (*9*). G can adopt at least three conformational states (*7*, *8*, *10*–*14*): the native prefusion state detected at the viral surface above pH 7; the activated hydrophobic state which interacts with the membrane as a first step of the fusion process (*11*); and the fusion-inactive postfusion conformation that is antigenically distinct from both the native and activated states. There is a pH-dependent equilibrium between the different states of G that is shifted toward the postfusion conformation at low pH (*15*). Thus, unlike fusogenic proteins from other viral families, the native prefusion conformation is not metastable (*9*). Indeed, the reversibility of the low pH induced conformational change is essential to allow G to be transported through the acidic compartments of the Golgi apparatus and to recover its native functional state at the viral surface (*16*).

We have recently determined the low-pH postfusion three-dimensional structure of the VSV G ectodomain (residues 1 to 422), generated by limited proteolysis of the virions with thermolysin (G_{th}) (*17*). In spite of having an unrecognized fold distinct from those of other fusion proteins previously described, the postfusion conformation of G displays the classic hairpin conformation of other viral fusogenic proteins [i.e., an elongated structure with the fusion domain and the TM domain at the same end of the molecule (*18*)]. As in class I fusion proteins (*19*–*21*), the postfusion trimer displays a six-helix bundle with the fusion domains at the N terminus of the central helices and the TM domains at the C terminus of the antiparallel outer helices. Each fusion domain bears two fusion loops located at the tip of an elongated β sheet, which is a marked convergence with class II fusion proteins (*22*–*24*). Unexpectedly, G turned out to be homologous to glycoprotein gB of herpesviruses, the atomic structure of which was published at the same time (*25*) despite a low pH induced conformational change of cellular viral G is reversible it remained unclear to what extent the pre- and postfusion conformations differed for this class of fusion proteins.

Among the different crystal forms obtained with G_{th} (*17*) (see also the materials and methods in the supporting online material), one of them, which was grown at pH 8.7, appeared to be particularly notable, because the asymmetric unit could not accommodate the postfusion form (125 Å in length) but was consistent with the presence of one protomer of the prefusion form (85 nm in length as measured for the RV G ectodomain by electron microscopy (EM) (*26*). This crystal structure of G_{th} was determined to 3.0 Å resolution by molecular replacement with the use of domains I, III, and IV (Table 1) of the low-pH form as search models. Data collection and refinement statistics are given in table S1. The structure of G_{th} is depicted in Fig. 1. Its length (88 Å), the location of the antigenic sites, and the comparison with the low-pH structure indicate that this G_{th} structure corresponds to the prefusion conformation of the molecule. The chain can be traced up to residue 413 (see the electron density for the final model in fig. S1). Clear density is also present for the first residues of both oligosaccharide chains (on N^{163} and N^{320}) (*27*), which were disordered in the structure of the low-pH form.

Table 1 Domain nomenclature used in the text. Root mean square deviation (RMSD) is between the pre- and postfusion structures. The number of alpha carbons (Cα) used in superposing the domains is indicated in parentheses.

Domain	Domain name	Color	Residues	RMSD
DI	Lateral domain	Red	1 to 17 and 310 to 382	0.42 Å (80 Cα)
DII	Trimerization domain	Blue	18 to 35, 259 to 309, and 383 to 405	-
DIII	PH domain	Orange	36 to 46 and 181 to 258	0.40 Å (82 Cα)
DIV	Fusion domain	Yellow	53 to 172	0.77 Å (94 Cα)
Cter	C-terminal part	Magenta	406 to 413	-
Rbl-II	Rigid block		1 to 25 and 273 to 382	0.56 Å (122 Cα)

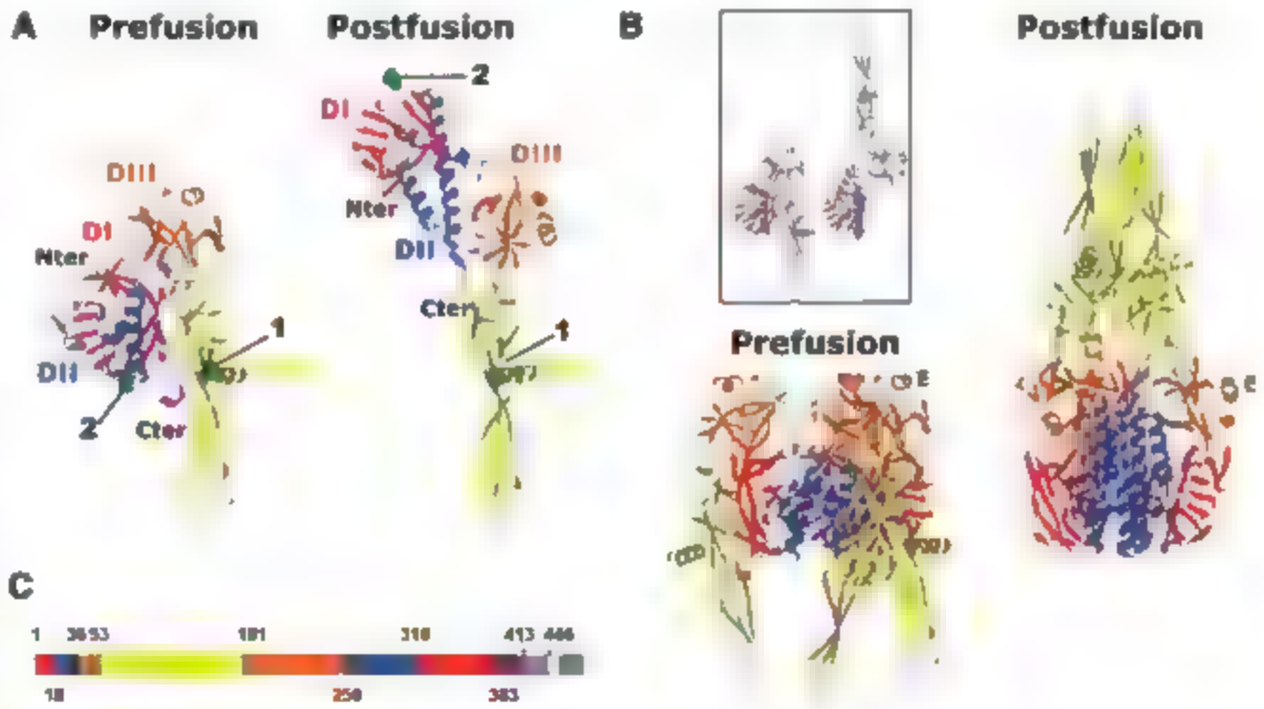


Fig. 1. Overall G_{th} structure in pre- and postfusion conformations. (**A**) View of the G protomers superimposed on their fusion domains (DIV) and colored by domain (as defined in Table 1) with the fusion loops in green. The two glycosylated asparagines [N^{163} (labeled "1") and N^{320} (labeled "2")] are displayed as dark green spheres. (**B**) View of the G trimers, colored by domains as in (A). The trimers were superimposed on the rigid blocks made of DI and the invariant part of DII (Rbl-II, defined in Table 1 and highlighted in the boxed inset for one protomer of each conformation). Helix E is indicated on both trimers. (**C**) Domain architecture of VSV G plotted on a linear diagram, color-coded according to Table 1 with domain boundaries numbered. The unobserved C-terminal segment is in gray with a checkerboard pattern for the TM domain. The regions that refold in the transition are hatched. All structural figures were generated with PyMOL (*38*).

The overall architecture of G_{fs} in its prefusion state resembles a tripod (Fig. 1B). Each leg is composed of a fusion domain with the fusion loops pointing toward the viral membrane. The last residues that we can see (including the conserved H^{407} and P^{408}) pack against the side of the fusion domain. This organization, which is reminiscent of the low-resolution structure of retroviruses' envelope spikes that was recently determined by EM (*28*, *29*), suggests that the TM segments are separate in the membrane. Nevertheless, we cannot exclude the possibility that the missing C-terminal segments of the ectodomain (residues 414 to 446) that lead to the TM segments come together toward the threefold axis.

In the tripod arrangement, the fusion domains are set wide apart, keeping the fusion loops separate (Fig. 1B, left). In contrast to class I and class II fusion proteins, the fusion loops of G are not buried at an oligomeric interface in the prefusion conformation. The hydrophobic residues Y^{116}, A^{117}, W^{72}, and Y^{73} are exposed (Fig. 1, A and B), even though they cluster near crystal contacts (fig. S3D). The tips of the fusion domains are the most flexible part of the structure (fig. S4) and thus are the least well defined in the electron density maps.

The conformational change involves a dramatic reorganization of the G molecule. Figure S2 shows a comparison of the secondary structure elements of the two conformations with their nomenclature. The pre- and postfusion states are related by flipping both the fusion domain and a C-terminal segment (composed of residues 383 to 413) relative to a rigid block (Rbl-II) made by the lateral domain and residues of the trimerization domain that include helix F2 of the prefusion form (Table 1 and Fig. 1B, inset). During the structural transition, both the fusion loops and the TM domain move ~160 Å from one end of the molecule to the other. Thus, the observed conformational change, although reversible, appears to be similar to that of paramyxovirus F glycoprotein (*30*). It also suggests that similar intermediates are formed during the fusion-associated refolding of G, HA, and paramyxovirus F glycoprotein (*19*, *30*). In one of these intermediates (Fig. 2C and movie S1), the fusion domain is projected at the top of the spike allowing the initial interaction with the target membrane.

In spite of large rearrangements in their relative orientation (Fig. 2, A and B), domains I, III, and IV retain their folded structure (Table 1 and Figs. 1A and 2). In this and the following paragraphs, we describe the conformational change of a protomer by considering Rbl-II as invariant (Fig. 1B, inset). The flippings of both the fusion domain and the TM segment relative to Rbl-II occur through a concerted rearrangement of distinct regions of the molecule. Although we have only snapshots of the initial and final states, analysis of the two structures (see the description of movie S1 in the supporting online material) suggests a plausible sequence of events leading from pre- to postfusion conformations.

The fusion domain is projected toward the target membrane through the combination of two movements (Fig. 2C): a 94° rotation around the hinge between the fusion and pleckstrin homology (PH) domains (Fig. 2A) and the repositioning of the latter domain at the top of the trimerization domain (Fig. 2B). The rotation involves the reorganization of two segments (residues 47 to 52 and 173 to 180) of the polypeptide chain. In the former segment, helix A^0 unfolds whereas, in the latter segment, helix C forms (Fig. 2A). Mutations $M^{44} \rightarrow$ V or I in RV G, which kinetically stabilize the native conformation (*31*), map to this region. Their location suggests that they impede the slight distortion of strands b and j of the PH domain that accompanies the movement.

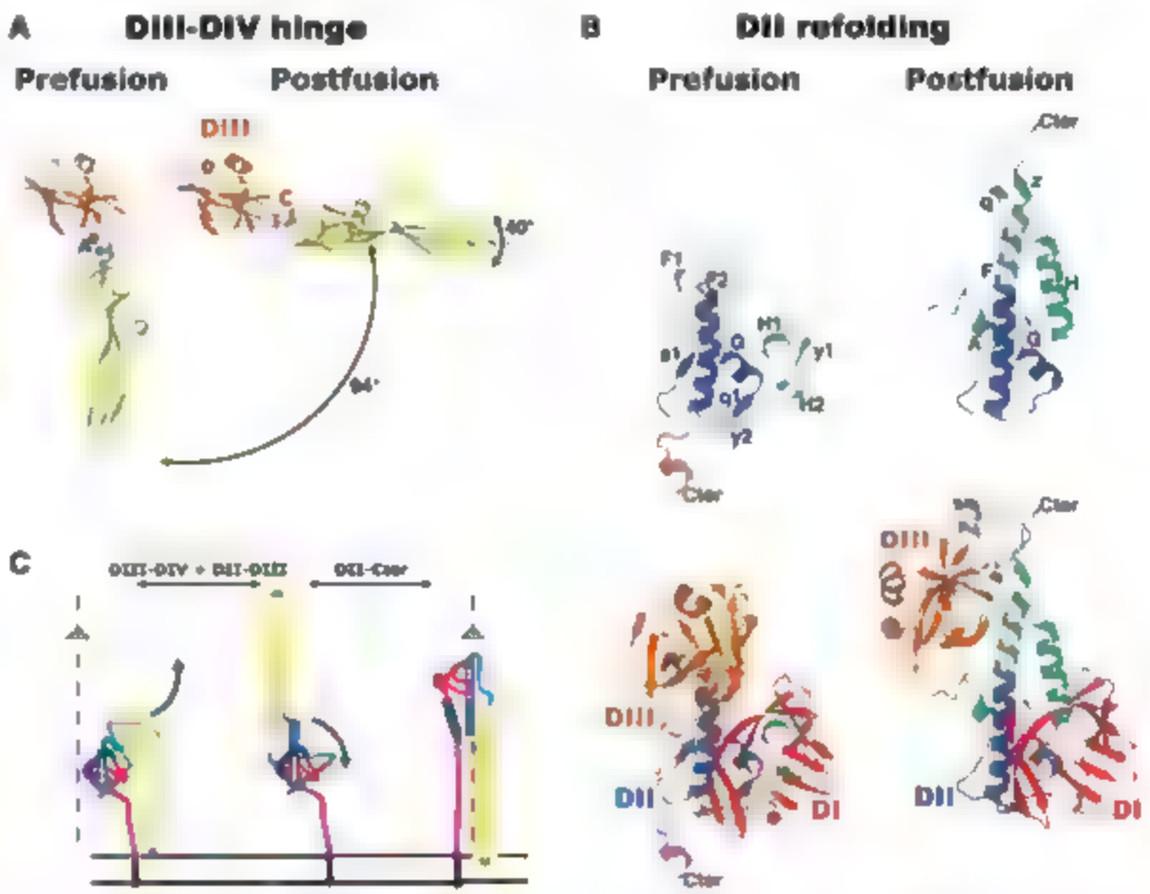


Fig. 2. Structural changes in the protomer between the pre- and postfusion conformations and relative movements of domains. In (A) and (B), fragments of the pre- and postfusion conformations are displayed to the left and right, respectively. Secondary structure elements of the prefusion form that refold are named and numbered according to fig. S2. (**A**) Relative movement of PH (DII, orange) and fusion (DIV, yellow) domains. The protomers are superimposed on DIII. Hinge residues 47 to 52 (prefusion helix A^0) and 173 to 180 (postfusion helix C) are colored in cyan and gray-blue, respectively. (**B**) Domain II refolding. DI and DIII are omitted in the top panels for clarity but are shown in the bottom panels to provide the relative orientations in the two forms. The protomers are superimposed on the invariant part of DII, which is indicated in dark blue, whereas the three segments that refold and/or relocate are indicated in shades of green. In the prefusion form, strands a^1 and y^1 form an interchain β sheet. The DIII-DIV hinge (bottom panels) is displayed and colored as in (A), with the two segments connected by a yellow bar to mark the location of the fusion domain. (**C**) Cartoon representation of the relative organization of domains with respect to the viral membrane during the conformational change. The one-sided black arrows indicate the relative movements of domains. The N- to C-terminal orientations of helices F2 (blue; left), F (blue; middle and right), and H (dark blue; right) are indicated with white arrows. Pre- (left) and postfusion (right) conformations are shown. The trimer axes are indicated. The middle cartoon shows how the fusion loops (in green) would be projected after the refolding of both the DIII-DIV hinge and the DII-DIII connection and before the C-terminal refolding of helix H.

The trimerization domain undergoes a major refolding event during the transition between the pre- and postfusion structures (Fig. 2B). This refolding drives the repositioning of the PH domain and the flipping of the C-terminal part and involves all three segments of the trimerization domain (Fig. 1C).

As a first step, central helix F2 (residues 276 to 294) is lengthened by the recruitment of a segment (made up of residues 263 to 275) to form the long helix F. The second segment that refolds is composed of residues 26 to 35, which, in the prefusion conformation, is buried in a groove of Rbl-II that is closed by residues 263 to 275. A sharp bend is introduced right after the conserved motif $C^{24}P^{25}$. The peptide bond between P^{25} and S^{26} flips, which redirects the polypeptide chain at an 80° angle, and short helix A (residues 24 to 29) is formed. The conformation of short strand a^1 (residues 22 to 24), involved in the interchain β-sheet a^1y^1 in the prefusion conformation, is unchanged although it is not paired to strand y^1 of the adjacent protomer in the postfusion conformation (Fig. 2B).

The small β-sheet q^1y^2 of the native form is then broken, although the individual strands q^1 and y^2 retain their β conformation in the postfusion form, and residues 384 to 400 (including helices H1 and H2 and strand y^1) refold into helix H. This helix then positions itself in the grooves of the central core in an antiparallel manner to form the six-helix bundle. This movement repositions the TM domains at the same end of the molecule as the fusion domains (Figs. 1B and 2B). Finally, residues 259 to 261 and 403 to 405, which are distant by ~30 Å in the prefusion conformation, form sheet qz that zips together helices F and H in the postfusion state (Fig. 2B). Strands q and z are already in an extended β structure in the native conformation primed to form sheet qz in the postfusion form.

The buried interface between two subunits in the trimer is 1600 $Å^2$ per protomer, as calculated by the Protein Interfaces, Surfaces, and Assemblies server (32). This value is less than half of that of the buried interface in the low-pH form. This explains the increased stability of the oligomeric structure of G at low pH (3). The interactions between protomers are all located in domain II (fig. S5) but are different from those observed in the postfusion form (Fig. 3, A and B). Not only is prefusion helix F2 shorter than postfusion helix F, it is also tilted and its C-terminal end is kept away from the trimer axis (Fig. 3A). This results from repulsive forces between the carboxylates of the three E^{286} amino acids (Fig. 3C). In contrast to the postfusion form, the main contribution to trimer stability is not due to the central helix bundle but appears to come from interchain β-sheet a^1y^1 [which must break during the fusogenic transition, before the formation of helix H (Fig. 2B)] and its environment, burying 1250 $Å^2$ per protomer (Fig. 3D). The conformational change occurs at the viral surface even in absence of the target membrane. This seems to be topologically impossible without transient dissociation of the trimer. This hypothesis is in agreement with the large differences in the trimeric interfaces between the native and the postfusion conformations of G.

A number of the few conserved residues (fig. S2) are involved in key networks of interactions that are different in the two forms (Fig. 4). This set of residues includes amino acids D^{137}, Y^{139}, H^{407} and P^{408} that cluster together in the postfusion conformation to stabilize β-sheet qz of the trimerization domain (Fig. 4B). In the prefusion conformation, the qz sheet does not exist. D^{137} and Y^{139} remain associated with the segment corresponding to the q strand and contribute to a network of hydrogen bonds that also involves conserved W^{236} of the PH domain (Fig. 4A, top). This network is disorganized during the rotation of the fusion domain relative to the PH domain (Fig. 2A). Conserved histidines: H^{407} involved in a salt bridge with D^{137} in the low-pH structure (Fig. 4B), H^{162} [previously shown to be involved in the interactions between fusion do-

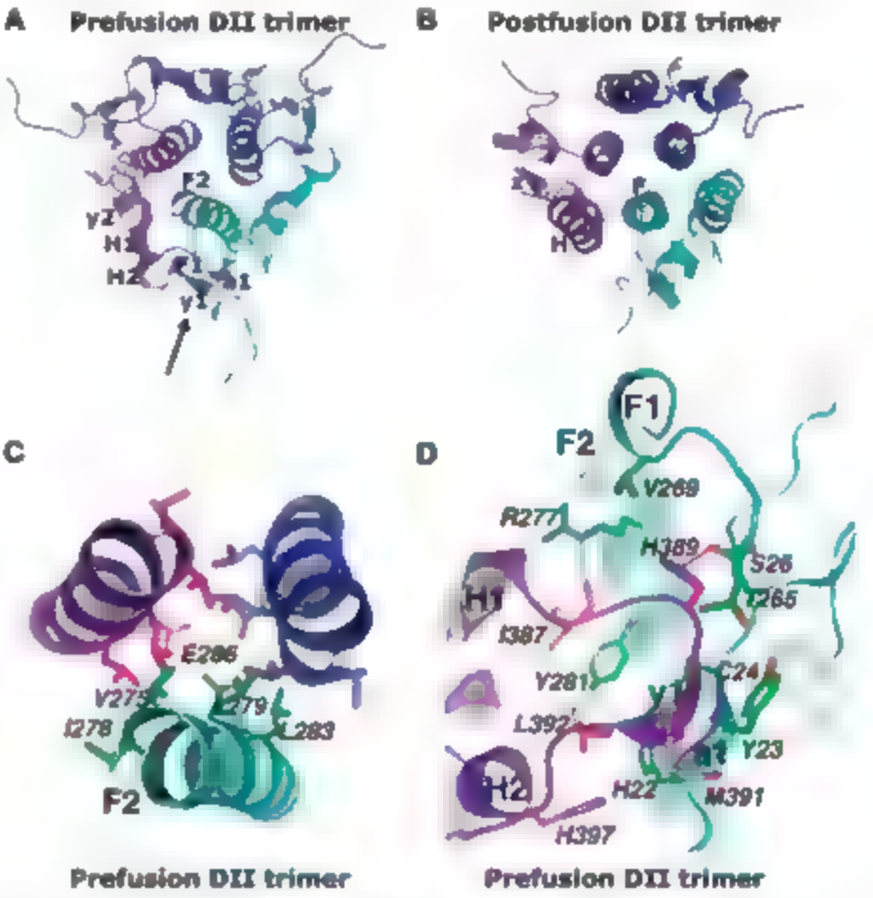


Fig. 3. The trimeric interface of the prefusion conformation [(A), (C), and (D)] as compared to that of the postfusion conformation (B). For clarity, only DII [the only domain involved in the interface in the prefusion conformation (fig. S5)] is represented, and the three protomers are colored in three shades of blue. Secondary structure elements that refold and/or relocate are labeled. (**A**) Top view (orientation as in fig. S5, looking down toward the viral membrane) of the trimeric interface of the prefusion conformation. The arrow indicates the viewpoint used in (D). (**B**) Trimeric interface of the postfusion conformation, superimposed on the invariant parts of DII in (A). The view therefore would now be from the membrane. (**C**) Zoom of image in (A) showing only the three helices F2 and the side chains involved in their interactions, which are colored by atom type (oxygen, red; nitrogen, blue; sulfur, yellow; carbon, green, magenta, or dark blue, depending on the protomer) and labeled. As in the postfusion state, V^{275} and T^{279} contribute to hydrophobic stabilizing interactions at the center of the molecule, but L^{283} now makes a lateral interaction with I^{278}. The three E^{286} amino acids in the center are 4 Å apart in native crystals. In the $YbCl_3$-derivative crystal used for refinement of the model (table S1), they chelate an ytterbium ion (not shown), bringing their side-chain oxygen atoms within 3.5 Å. (**D**) Close-up view of the outer region of the prefusion trimeric interface seen from the side. Contact residues are colored as in (C), with main-chain atoms included only when they participate in the contacts. Besides the canonical hydrogen bonds of the β sheet, the interface is stabilized through extended van der Waals contacts and a hydrogen bond between the imidazole ring and the carboxyl group of T^{265} of the neighboring protomer. Finally, carboxyl groups of L^{384} and I^{387} make two hydrogen bonds with the guanidium group of R^{277} of the other chain. These three hydrogen bonds are displayed as magenta dashed lines.

mains in the low-pH conformation (*17*)], and H^{60} cluster together (H^{60} is absent in RV G but H^{88}, which corresponds to S^{84} in VSV G replaces it) (Fig. 4A, bottom). Protonation of these residues at low pH is likely to destabilize the interaction between the C-terminal segment of G_{ch} and the fusion domain in the prefusion conformation, priming the initial movement of the fusion domain toward the target membrane. Conversely, the acidic amino acids that were either buried at the trimer interface (D^{268}) or brought close together (D^{274} with D^{395} and E^{276} with D^{395}) in the postfusion acidic conformation (*17*) are solvent-exposed in the prefusion state (not shown). Thus, the histidines in the prefusion form and the acidic residues in the postfusion form appear to constitute two pH-sensitive molecular switches.

The major antigenic sites of rhabdoviruses are located in the lateral and PII domains (*4*–*6*) (Fig. 5). The accessibility of antigenic sites to antibodies has been studied in detail for RV G. Antibodies directed against RV G site II are unable to recognize the protein in its low-pH conformation (*7*, *15*). Indeed, during the structural transition, this site moves from the top of the molecule to a less accessible location at the surface of the virus. Conversely, the N-terminal epitope of RV G (NS) is only accessible in the low-pH conformation at the viral surface (*37*). Finally, RV G minor site a is recognized in both conformations (*7*). As for monoclonal antibody 17D2 (*34*) that binds only the prefusion conformation (*34*), its epitope is located in the segment of helix F that is unfolded in the native structure.

The cellular receptor of VSV G has not been identified. Nevertheless, a canyon located between the lateral and PII domains is exposed at the top of the molecule and could be involved in ligand binding (Fig. 5A). In support of this, residues 330 and 333 of RV G, which are involved in the recognition of the putative viral receptor p75 (low-affinity nerve growth factor receptor) (*35*) and which affect viral pathogenesis, align with residues 331 and 334 of VSV G, which are located at either end of the canyon.

In a previous study, we estimated the minimal number of trimeric spikes involved in the formation of a RV fusion complex as about 15 (*15*). At the viral surface, a local organization of the spikes resembling the P6 lattice found in the crystal (in which all the spikes are oriented identically, with the major antigenic sites exposed at their tops) (fig. S3) might organize the glycoproteins in an optimal manner for a concerted conformational change. It might also facilitate the formation of the initial intermediates on the fusion pathway. Indeed, the initial lipid deformations leading to the formation of the stalk and the initial fusion pore (*36*) can form inside the inner rim of such a hexagon. Reinforcing the idea that the P6 organization may reflect the structure of a fusion relevant complex, a local hexagonal lattice of spikes of similar dimensions has been observed at low temperature under mildly acidic conditions at the surface of some RV G mutants that were affected in the kinetics of their low pH induced structural transition (*31*).

It is often considered that fusogenic proteins drive membrane fusion by coupling irreversible protein refolding to membrane deformation (*37*). At least for rhabdoviral G, this is not the case. Rather, it appears that a concerted cooperative change of a large number of glycoproteins (perhaps organized in a hexagonal lattice, like the one present in the crystals) is used to overcome the high energetic barrier encountered during fusion.

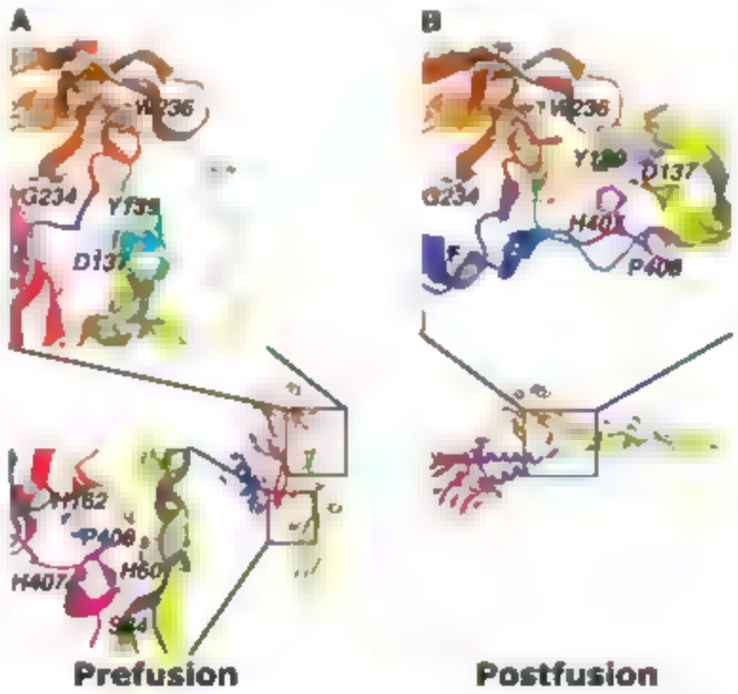


Fig. 4. Alternative networks of conserved residues in the pre- (**A**) and postfusion (**B**) conformations. The orientation is as that in Fig. 2A. Conserved residues are displayed in stick representation (main-chain atoms are not shown unless they participate in interdomain contacts). Hydrogen bonds are displayed as magenta dashed lines. [(A) and (B), top] Close-up views of the DIII-DIV connection are shown. The prefusion hydrogen bonds of Y^{133} to the main chain of W^{234} are relocated to the postfusion αz sheet, whereas D^{137} switches from making a bidentate hydrogen bond to the main chain to engaging in a salt bridge with H^{407} [(A), bottom]. A close-up view of the prefusion DIV-Cter interface that has to be disrupted for DIV to move is shown. Note the cluster of conserved histidines, including H^{407}.

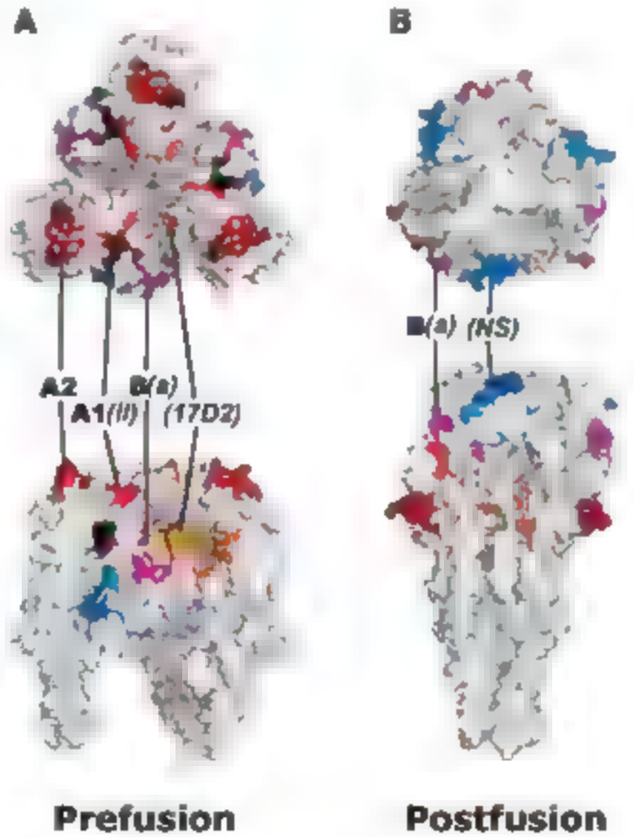


Fig. 5. Antigenic sites of Rhabdoviridae mapped onto the surface of the pre- (**A**) and postfusion (**B**) VSV G trimers. Sites are colored on both forms and labeled on the form(s) in which they are recognized. VSV sites are labeled in bold, and RV sites are labeled in italics within parentheses. VSV sites A1 (residues 37 to 38, corresponding to RV antigenic site II located on segments composed of residues 34 to 42 and 198 to 200) and A2 (located at the surface of helix E indicated in Fig. 1) are indicated in shades of red. The RV G site recognized by antibody 17D2 (between residues 255 and 270) is in orange. NS (extending from amino acid 10 to 15) is in dark blue. VSV site B (extending from amino acid 341 to 347), corresponding to RV G minor antigenic site a (amino acid 340 to 342), is in magenta. In the prefusion conformation, the cleft between DI and DIII is colored black. It is flanked by residues 331 and 334, in gray, whose counterparts in RV affect virulence.

References and Notes

1. J. K. Rose, M. A. Whitt, in *Fields' Virology* D. M. Knipe, P. M. Howley, Eds. (Lippincott, Williams & Wilkins, Philadelphia, ed. 4, 2001), pp. 1221–1244
2. B. L. Rao *et al.*, *Lancet* **364**, 869 (2004)
3. I. Le Blanc *et al.*, *Nat. Cell Biol.* **7**, 653 (2005).
4. L. H. Lee, Y. Li, R. M. Snyder, R. R. Wagner, *Virology* **163**, 341 (1988)
5. A. Benmansour *et al.*, *J. Virol.* **65**, 4198 (1991).
6. S. B. Vandepol, L. Lefrancois, J. J. Holland, *Virology* **148**, 312 (1986).
7. Y. Gaudin, C. Tuffereau, D. Segretain, M. Knossow, A. Flamand, *J. Virol.* **65**, 4853 (1991)
8. R. W. Doms, D. S. Keller, A. Helenius, W. E. Balch, *J. Cell Biol.* **105**, 1957 (1987)
9. Y. Gaudin, *Subcell. Biochem.* **34**, 379 (2000)
10. Y. Gaudin, R. W. Ruigrok, M. Knossow, A. Flamand, *J. Virol.* **67**, 1365 (1993)
11. P. Durrer, Y. Gaudin, R. W. Ruigrok, R. Graf, J. Brunner, *J. Biol. Chem.* **270**, 17575 (1995)
12. B. L. Fredericksen, M. A. Whitt, *Virology* **217**, 49 (1996)
13. C. C. Pak, A. Puri, R. Blumenthal, *Biochemistry* **36**, 8890 (1997)
14. F. A. Carneiro, A. S. Ferradosa, A. T. Da Poian, *J. Biol. Chem.* **276**, 62 (2001)
15. S. Roche, Y. Gaudin, *Virology* **297**, 128 (2002)
16. Y. Gaudin, C. Tuffereau, P. Durrer, A. Flamand, R. W. Ruigrok, *J. Virol.* **69**, 5528 (1995).
17. S. Roche, S. Bressanelli, F. A. Rey, Y. Gaudin, *Science* **313**, 187 (2006).
18. M. Kielian, F. A. Rey, *Nat. Rev. Microbiol.* **4**, 67 (2006)
19. P. A. Bullough, F. M. Hughson, J. J. Skehel, D. C. Wiley, *Nature* **371**, 37 (1994)
20. J. J. Skehel, D. C. Wiley, *Cell* **95**, 871 (1998)
21. H. S. Yin, R. G. Paterson, X. Wen, R. A. Lamb, T. S. Jardetzky, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9288 (2005).
22. Y. Modis, S. Ogata, D. Clements, S. C. Harrison, *Nature* **427**, 313 (2004)
23. S. Bressanelli *et al.*, *EMBO J.* **23**, 728 (2004).
24. D. L. Gibbons *et al.*, *Nature* **427**, 320 (2004).
25. E. E. Heldwein *et al.*, *Science* **313**, 217 (2006).
26. Y. Gaudin, R. W. Ruigrok, C. Tuffereau, M. Knossow, A. Flamand, *Virology* **187**, 627 (1992)
27. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
28. F. Forster, O. Medalia, N. Zauberman, W. Baumeister, D. Fass, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4729 (2005)
29. P. Zhu *et al.*, *Nature* **441**, 847 (2006)
30. H. S. Yin, X. Wen, R. G. Paterson, R. A. Lamb, T. S. Jardetzky, *Nature* **439**, 38 (2006)
31. Y. Gaudin, H. Raux, A. Flamand, R. W. Ruigrok, *J. Virol.* **70**, 7371 (1996)
32. E. Krissinel, K. Henrick, in *Comp.Life 2005* M. R. Berthold, R. Glen, K. Diederichs, O. Kohlbacher, I. Fischer Eds., *Lecture Notes in Bioinformatics*, vol. 3695 (Springer-Verlag, Berlin, 2005), pp. 163–174
33. F. Lalay, A. Benmansour, K. Chebli, A. Flamand, *J. Gen. Virol.* **77**, 339 (1996)
34. Y. Gaudin, *J. Virol.* **71**, 3742 (1997)
35. C. Tuffereau, J. Benejean, D. Blondel, B. Kiefer, A. Flamand, *EMBO J.* **17**, 7250 (1998)
36. L. V. Chernomordik, M. M. Kozlov, *Cell* **123**, 375 (2005)
37. C. M. Carr, C. Chaudhry, P. S. Kim, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14306 (1997)
38. W. L. DeLano, The PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA, 2002), available at www.pymol.org

39. We thank A. Flamand for constant support on this project; J. Lepault, R. Ruigrok, M. Knossow, A. Benmansour, C. Tuffereau, and D. Blondel for helpful discussions at different stages of this work, and C. Maheu for virus purification. Data collections were performed at the Swiss Light Source (SLS), Paul Scherrer Institut, Villigen, Switzerland, and at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We acknowledge the help of T. Tomizaki (beamline X06SA, SLS), G. Leonard and D. Bourgeois (beamlines ID29 and ID23-2, ESRF), and S. Duquerroy and G. Squires in data collection. We acknowledge support from the CNRS and INRA, the CNRS program "Physique et Chimie du Vivant," the INRA Animal health department program "Les virus des animaux et leurs interactions avec la cellule," the Ministère de l'éducation nationale, de la recherche et de la technologie program "Action Concertée Incitative blanche," and the Agence Nationale de la Recherche program. S.R. was the recipient of an Agence Nationale de Recherche sur le Sida fellowship during part of this project. Coordinates and structure factors have been deposited with the Protein Data Bank under accession code 2j6j.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5813/843/DC1
Materials and Methods
Figs. S1 to S5
Table S1
References
Movie S1

29 September 2006; accepted 3 January 2007
10.1126/science.1135710

Relative Impact of Nucleotide and Copy Number Variation on Gene Expression Phenotypes

Barbara E. Stranger,[1] Matthew S. Forrest,[2] Mark Dunning,[2] Catherine E. Ingle,[1] Claude Beazley,[3] Natalie Thorne,[2] Richard Redon,[2] Christine P. Bird,[2] Anna de Grassi,[3] Charles Lee,[4,5] Chris Tyler-Smith,[1] Nigel Carter,[1] Stephen W. Scherer,[6,7] Simon Tavaré,[2,8] Panagiotis Deloukas,[1] Matthew E. Hurles,[1]* Emmanouil T. Dermitzakis[1]*

Extensive studies are currently being performed to associate disease susceptibility with one form of genetic variation, namely single-nucleotide polymorphisms (SNPs). In recent years, another type of common genetic variation has been characterized, namely structural variation including copy number variants (CNVs). To determine the overall contribution of CNVs to complex phenotypes, we have performed association analyses of expression levels of 14,925 transcripts with SNPs and CNVs in individuals who are part of the International HapMap project. SNPs and CNVs captured 83.6% and 17.7% of the total detected genetic variation in gene expression, respectively, but the signals from the two types of variation had little overlap. Interrogation of the genome for both types of variants may be an effective way to elucidate the causes of complex phenotypes and disease in humans.

Understanding the genetic basis of phenotypic variation in natural populations is currently one of the major goals in human genetics. Gene expression (the transcription of DNA into mRNA) has been interrogated in a variety of species and experimental scenarios in order to investigate the genetic basis of variation in gene regulation (*1*–*8*), as well as to tease apart regulatory networks (*9*, *10*). In some respects, a comprehensive survey of gene expression phenotypes (steady state levels of mRNA) serves as a proxy for the breadth and nature of phenotypic variation in human populations (*11*). Much of the observed variation in mRNA transcript levels may be compensated at higher stages of regulatory networks, but an understanding of the nature of genetic variants that affect gene expression will provide an essential framework and model for elucidating the causes of other types of phenotypic variation. Single-nucleotide polymorphisms (SNPs) have long been known to be associated with phenotypic variation either through direct causal effects or by serving as proxies for other causal variants with which they are highly correlated (i.e., in linkage disequilibrium) (*1*, *2*, *12*). An understanding of this association has been facilitated by the validation of millions of SNPs by the International HapMap project (*13*). However, during the last few years, structural variants, such as copy number variants (CNVs)—defined as DNA segments that are 1 kb or larger in size present at variable copy number in comparison with a reference genome (*14*)—have attracted much attention (*2*). It has become apparent that they are quite common in the human genome (*15*–*19*) and can have dramatic phenotypic consequences as a result of altering gene dosage, disrupting coding se-

[1]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK. [2]Department of Oncology, University of Cambridge, Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK. [3]Istituto di Tecnologie Biomediche-Sezione di Bari, Consiglio Nazionale delle Ricerche (CNR), 70126 Bari, Italy. [4]Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA. [5]Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA 02142, USA. [6]The Centre for Applied Genomics and Program in Genetics and Genomic Biology, The Hospital for Sick Children, MaRS Centre, Toronto, Ontario, M5G 1L7, Canada. [7]Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario, Canada. [8]Program in Molecular and Computational Biology, University of Southern California, Los Angeles, CA 90089-2910, USA.

*To whom correspondence should be addressed. E-mail: md4@sanger.ac.uk (E.T.D.); meh@sanger.ac.uk (M.E.H.)

quences, or perturbing long-range gene regulation (*20*, *21*). Evidence has been presented that increased copy number can be positively (*18*, *22*) or negatively (*23*) correlated with gene expression levels (for example, deletion of a transcriptional repressor could serve to elevate gene expression) but the relative contribution of such large genetic variants (i.e., CNVs) and smaller variants (i.e., SNPs) to phenotypic variation has not been evaluated. It is also still unknown whether SNPs can serve as proxies to CNVs (*24*, *25*) and whether the complex nature of some CNVs requires that they be surveyed directly (*26*). We have used the phase I HapMap SNPs (*13*) and the recently described CNV data ascertained in the same HapMap populations (*26*) for correlation with genome-wide gene expression variation in the same individuals.

Gene expression was interrogated in lymphoblastoid cell lines of all 210 unrelated HapMap individuals (*13*) from four populations (CEU: 60 Utah residents with ancestry from northern and western Europe; CHB: 45 Han Chinese in Beijing; JPT: 45 Japanese in Tokyo; YRI: 60 Yoruba in Ibadan, Nigeria) in four technical replicates (see Methods). Out of the 47,294 transcripts that were interrogated, the normalized values for 14,925 transcripts (14,072 genes) were included in the analysis (see Methods and (*27*)). The SNP genotypes from phase I HapMap (*28*) were used in the analysis (see Methods). CNV data were represented by $\log_2$ ratios from comparative genomic hybridization (CGH) of each HapMap individual against a common reference individual on an array comprising 26,574 large-insert clones covering 93.7% of the euchromatic portion of the genome (*26*, *29*). $\log_2$ ratios from two sets of clones were analyzed: the whole set of 24,963 autosomal clones (CGH clones) and the 1322 autosomal clones corresponding to CNVs present in at least two HapMap individuals (CNV clones) (*26*). We excluded genes on sex chromosomes because of their imbalance in males and females. We performed linear regression (on each of the four populations separately) between normalized quantitative gene expression values and SNP genotypes or clone $\log_2$ ratios that were near the gene (SNP position or clone midpoint within 1 Mb and 2 Mb, respectively, of the probe midpoint position). We used different window sizes for SNPs and clones because clones are large (median size of ~170 kb) and structural variants can exert long-range effects (*21*), so a 2-Mb window is more appropriate. Statistical significance was evaluated through the use of permutations (*30*), as previously described (*1*), and a corrected P value threshold of 0.001 was applied (see Methods). Repeated permutation exercises showed that our permutation thresholds were very stable (see table S1). We tested a large number of genes so an additional correction was required. This could be done either by adjusting the threshold to a new corrected threshold above which all genes are expected to be significant (e.g., Bonferroni correction) or by setting the threshold to a value that generates a satisfactory false-discovery rate (FDR). We have used the second, and we have estimated the FDR on the basis of the number of genes tested and have required that, in all cases, at least 80% of the genes called significant are estimated to be truly significant. Given that there are 14,072 genes that lie within 1 Mb of SNPs and within 2 Mb of the full set of CGH clones, and ~7150 genes that lie within 2 Mb from the CNV clones (from 7135 to 7191 depending on the population, owing to missing data), we expect this analysis to generate false-positive association signals for approximately 14 and 7 genes, respectively, in each population.

Of the 14,072 genes tested, we detected significant associations with at least one SNP for 323, 348, 370, and 411 genes for CEU, CHB, JPT, and YRI, respectively (e.g., Table 1 and table S1). These comprise a total of 888 non-redundant genes of which 331 (37%) were replicated at the same significance level in at least one other population, and of those, 67 (8%) were significant in all four populations (Table 2 and table S2). As expected, we have limited power to detect weak effects because of the small sample sizes. The minimum detected squared regression coefficient (r^2) which reflects the proportion of expression variance accounted for by the linear association with allele counts was 0.27. However, some very strong effects were detected that, in some cases, had an r^2 close to 1 (Fig. 1 and fig. S1). We detected a strong preference for associated SNPs to be close to their respective genes, most of which were within 100 kb of the interrogated expression probe (Fig. 1, A and C). In summary, we detected a large number of regions that appear to carry genetic variation affecting gene expression. To evaluate the effect of experimental variation and, hence, the robustness of our associations, we compared the list of gene expression associations from our previous study (*1*) in which we detected 63 expression associations significant at the 0.05 permutation threshold in the CEL population. Of those 63 expression phenotypes, 47 went into the current analysis, of which 43 (91.5%) were called significant at the same permutation threshold (0.05) in the same population. The previous study was performed with different batches of cells, by using RNA extracted in a different laboratory, with RNA levels quantified on a different type of array (custom versus genome-wide array), so the high degree of experimental and statistical replication strongly suggests that the signals we detected are robust and stable to experimental variation in expression measurements.

Of the 14,072 genes tested, we detected significant associations with at least one of the 24,962 autosomal CGH clones in 85, 44, 58, and 96 genes in CEU, CHB, JPT, and YRI, respectively (238 nonredundant genes), of which 28 (12%) were replicated at the same significance level in at least one other population, and of those, 5 (2%) were significant in all four populations (Fig. 2, Table 1, and table S3 and figs. S2 and S3). Not all associated clones were within CNVs defined using the stringent criteria of (*26*) [119 out of 303 (39%) associated clones were previously defined as CNVs], and it is likely that some of these clones encompass smaller CNVs that are detectable though asso-

Table 1. Numbers of genes with significant associations to SNPs (SNP-probe distance < 1 Mb), all CGH clones (clone-probe distance < 2 Mb), or CNV clones (clone-probe distance < 2 Mb) as assessed by permutations, together with the numbers of overlaps between SNP-associated genes and CGH or CNV clone–associated genes (probe-variant distance < 1 Mb for both SNPs and clones) (see table S4).

Gene population	CNV (2 Mb)		SNP	CNV (1 Mb) + SNP overlap	
	CGH clones	CNV clones		CGH clones	CNV clones
			Permutation threshold 0.01		
CEU	362	138	643	14	15
CHB	221	110	673	10	9
JPT	319	134	752	13	14
YRI	481	166	815	14	11
Nonredundant	1246	451	1886	28	16
			Permutation threshold 0.001		
CEU	85	40	323	9	8
CHB	44	32	348	5	6
JPT	58	40	370	8	6
YRI	96	42	411	7	6
Nonredundant	238	99	888	15	12
			Permutation threshold 0.0001		
CEU	32	18	198	5	6
CHB	14	19	204	4	4
JPT	23	20	217	6	5
YRI	27	16	251	2	2
Nonredundant	69	39	526	8	8

cations of $\log_2$ ratios across a population, but cannot be detected as extreme outliers in their $\log_2$ ratios in any one individual [as is required for classification as a CNV in (*26*)] (see example below). For 36 common (minor allele frequency > 0.05) CNVs (encompassing 99 CGH clones), accurate CNV genotypes were available. We used these genotypes to validate the statistical power of performing association analysis using $\log_2$ ratios directly rather than genotypes. There was strong correlation between r^2 values or P values generated using the $\log_2$ ratio signals or the CNV genotypes (Pearson correlation coefficients > 0.9), indicating that $\log_2$ ratios can be used directly.

Little prior data exists on CNV-expression associations against which to compare and demonstrate the robustness of our associations. One recent study (*18*) demonstrated three associations between common deletions and gene expression in a subset of the CEU. Two of these deletions are covered by our CGH data. The reported expression association caused by the largest of these two deletions is also captured in our analysis (influencing *UGT2B17*), and we extend this observation to show that this deletion also affects the expression of three other nearby genes (*UGT2B7*, *UGT2B10*, and *UGT2B11*) and that these associations replicate across all four populations. The smaller deletion of only 18 kb, reported previously (*18*) as affecting expression of *GSTM1*, is below the expected resolution of the CGH data. Nonetheless, we observe an association that, although it does not pass our stringent permutation threshold (0.001), has significant nominal P values in all four populations ($P_{CEU} = 0.0292$; $P_{YRI} = 0.0018$; $P_{JPT} = 0.0408$; $P_{CHB} = 0.0185$). Thus, suggests that effects of CNVs far smaller than genomic regions that met our criteria to be called a CNV within the CGH platform can be detected and replicated in multiple populations with our analysis.

Having investigated the potential contribution of CNV to variation in gene expression by using data from all CGH clones, we interrogated the nature of CNV effects on gene expression in finer detail by performing association tests of 1322 clones within high confidence CNVs (see above) with expression of the 14,072 genes, in order to generate a set of high stringency associations for which the presence of an underlying CNV has already been validated. Significant associations with at least one of the 1322 CNV clones were detected for 40, 32, 40, and 42 genes in CEU, CHB, JPT, and YRI, respectively (99 nonredundant genes) (table S4). Thirty four of the 99 genes (34%) associated with CNV clones have a significant signal in at least two populations (Table 2), of which 7 (7%) were associated in all populations. Some CNV clones were associated with more than one gene in the same population; a notable example was a single CNV clone associated with expression of four genes in all populations (*UGT2B* genes, see above). CNVs detected by CGH can be classified into five classes: deletion, duplication, deletion and duplication at the same locus, multiallelic, and complex (*26*). We find all classes of CNV represented among the significant associations. Despite the clear preference for genes to lie close to their associated CNVs (Fig. 1, B and D), 53% of the expression probes associated with a CGH clone were located outside the CNVs encompassing that clone (*26*). This suggests that rather than altering gene dosage, about half the CNV effects are caused by disruption of the gene (some parts of the gene but not the probe are within in the CNV) or affect regulatory regions and other functional regions that have an impact on gene expression. When we extended our analysis to consider associations between genes and CNVs up to 6 Mb apart, we detected a few significant long-distance associations beyond 2 Mb (table S5). These types of long-range effects are becoming more apparent through recent studies looking in detail at specific genomic regions (*20*, *31*). A small minority (5 to 15%) of the significant CNV-expression associations have a negative correlation between copy number and gene expression, which suggests that not all the detected effects are of the conventional type, wherein gene expression levels increase with gene copy number (table S3). Almost all (32 out of 34) of the associations that are shared between populations also exhibit the same direction of correlation in all populations. The two exceptions could result from the CNVs being in linkage disequilibrium with different regulatory variants in different populations or because of SNP × CNV interactions. However, the strong bias toward positive correlations between copy number and expression levels implies that the vast majority of these associations are attributable to the CNV itself, and not to a linked variant.

We next determined whether the same associations were also captured by SNPs (Fig. 2 and figs. S3 and S4). We only considered those CGH clones or CNVs within 1 Mb of the probe so that the analysis is comparable to that of the SNPs (total of 188 and 84 genes for CGH clones and CNVs, respectively). We expect some of the CNVs to be correlated with SNPs via common genealogical history (linkage disequilibrium) and therefore their effect on gene expression would also be captured by SNP associations. Fewer than 20% (in all populations) of the detected CGH clone associations overlapped with SNP associations (Table 1), even when we included CGH and SNP associations with the same gene but in different populations [28 out of 188 (14%) genes with significant CGH clone associations also had a SNP association in any population]. The same is true of CNV clone associations. Only 15 of 84 genes (18%) with CNV clone associations within 1 Mb also had a SNP association in any population, and if we required the association in the same population only 12 (14%) of genes had a SNP association. On the basis of previous work (*26*, *32*) on patterns of linkage disequilibrium around CNVs (*26*), we considered that this low overlap between CNV or CGH clone associations with SNP associations might be due in part either to a low density of successfully genotyped SNPs around some CNVs or to the suppression of apparent LD by recurrent mutation at some CNVs. Segmental duplications (SDs) are the primary cause of low

Table 2. Sharing of associations between populations.

	CGH clone (2 Mb)	CNV clone (2 Mb)	SNP (1 Mb)
CEU-CHB-JPT-YRI	5	7	67
CEU-CHB-JPT	2	4	48
CEU-CHB-YRI	1	0	11
CEU-JPT-YRI	1	0	12
CHB-JPT-YRI	3	3	28
CEU-CHB	1	3	18
CEU-JPT	2	0	15
CEU-YRI	6	6	36
CHB-JPT	4	5	51
CHB-YRI	1	3	18
JPT-YRI	2	3	27
CEU only	67	20	116
CHB only	27	7	107
JPT only	39	18	122
YRI only	77	20	212
Sum	238	99	888
Gene associations in at least two populations	28	34	331
Percentage of total	0.12	0.34	0.37
Gene associations in single populations	210	65	557
Percentage of total	0.88	0.66	0.63

SNP densities in HapMap Phase I because of the difficulties in developing robust SNP genotyping assays within them (*13*). We did not observe enrichment of segmentally duplicated sequences within the CGH and CNV clones that did not share signals with SNPs relative to those CGH and CNV clones that did share signals with SNPs. However, we observe a 2.5-fold excess of compound CNVs [CNVs with more than one mutation event, on the basis of classification of the CNVs in (*26*)] in associations that are not shared with SNPs relative to those that are shared (Fisher's exact test $P < 0.001$). Thus our analysis suggests that recurrent mutation is a key factor reducing overlap between CNV and SNP associations.

CNV associations that were also detected with SNPs were clearly biased toward large effect sizes (tables S1 and S3). Of the 12 genes with both SNP and CNV associations in the same population, 8 shared the association in two or more populations (giving a redundant total across the four populations of 26 shared CNV and SNP associations). The ratio of 8 out of 12 (67%) population shared associations is larger than that observed in all CNV associations (34 out of 99 = 34%) potentially suggesting that associations with higher frequency older CNVs are more likely to be captured by SNPs. For the 26 associations (representing 23 genes [see above]) captured both by CNVs and SNPs in the same population, we observed that SNPs and CNVs were themselves highly correlated for 23 out of 26 SNP-CNV pairs (Pearson correlation, $P < 0.001$) suggesting that for these cases the CNV and SNP captured the same effect, and that only a small fraction of the associations captured both by SNPs and CNVs occurs by chance. In summary, 87 out of 99 (87%) of genes with a significant CNV association are not associated with SNPs.

The large-scale (typically > 100 kb) copy number variation analyzed here appears to be associated with about 10 to 25% as many gene expression phenotypes as captured by ~700,000 SNPs, and the majority of these effects cannot be explained by altered dosage of the entire gene but by gene disruption and its impact on the regulatory landscape of the region where these CNVs occur. When we restrict the analysis to within 1 Mb of the probe of the expressed gene we detected 1061 genes associated with CGH clones or SNPs, 17.7% of which are associated with CGH clones, 83.6% with SNPs, and 1.3% with both. Of the 972 genes associated with CNV clones or SNPs, 8.75% are associated with CNV clones, 92.5% with SNPs, and 1.25% with both. Whereas the phase I HapMap SNPs likely capture a large fraction of the SNP effects in the genome (*13*), only a small minority of the CNVs in the genome were considered here. CNVs < 100 kb in length are far more numerous than CNVs >100 kb in length (*19*). As a consequence, 8.75 to 17.7% is a minimal estimate of the proportion of heritable gene expression variation that is explained by copy number variation.

Our study has attempted to evaluate the relative impact of CNVs and SNPs on phenotypic variation in human populations. Within the limitations of our samples, tissue type, SNP coverage, and CNV resolution, each type of genetic variation captures a substantial number of largely mutually exclusive effects on gene expression. We also demonstrate that both CNV and

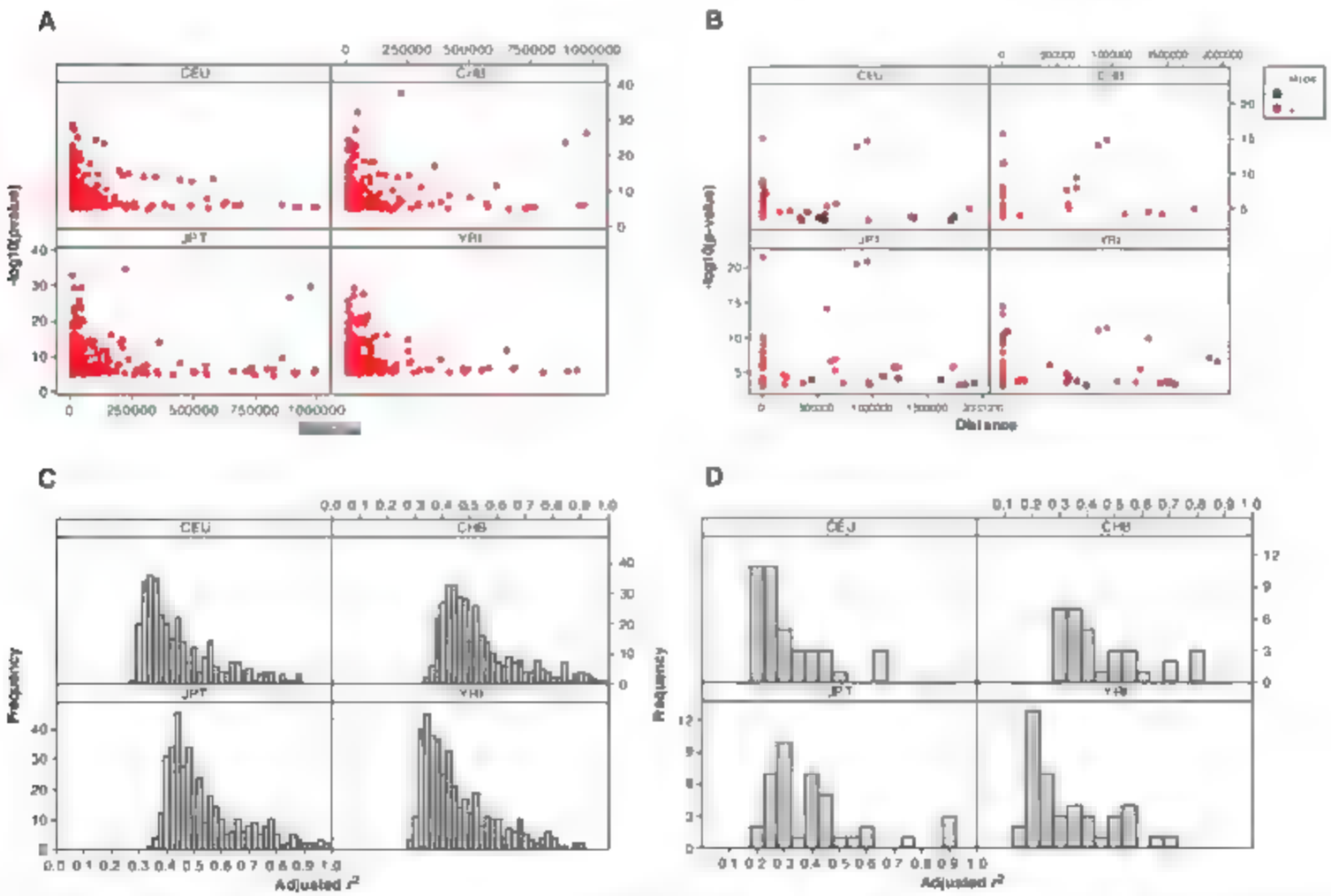


Fig. 1. Strength of association as a function of distance between (**A**) SNP and probe and (**B**) CNV and probe. Positive associations between mRNA levels and clone $\log_2$ ratios are shown in red, negative associations in black. Distance equal to zero corresponds to the probe residing within the CNV. In each population panel, only the details for the most significant association per significant gene are shown. Distribution of r^2 values for the most significant association per significant gene for (**C**) SNP-expression associations and (**D**) clone-expression associations.

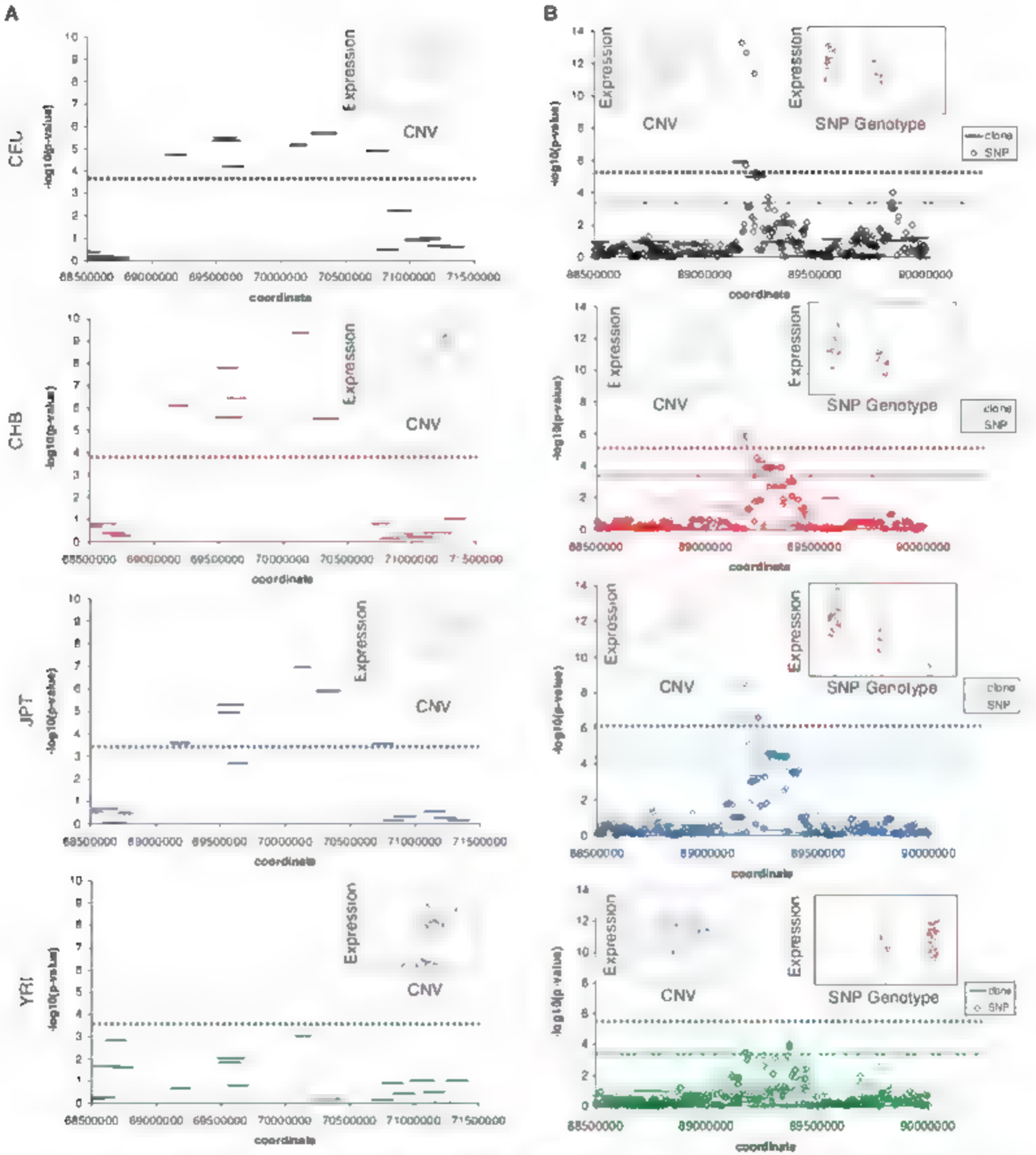


Fig. 2. Examples of SNP-expression and clone-expression associations in the four HapMap populations. (**A**) Clone-expression association for *SMN2*, chromosome 5 (chr 5). Significant associations between clones and expression are observed in CEU, CHB, and JPT, but not in YRI. (**B**) SNP-expression and clone-expression association for *GBP3*, chr 1. Both SNPs and clones are significantly associated with expression of *GBP3* in CEU, CHB, and JPT, but not in YRI. In each plot, dotted lines show the 0.001 permutation significance threshold. For clone–expression associations, all clones in the window are shown; however, the significance threshold was determined by permuting data only from those clones in CNVs where the CNV was present in at least two HapMap individuals. All coordinates shown are from Build 35 of the human genome. Inset panels show the relation between mRNA levels and SNP genotypes or clone $\log_2$ ratios, for the most significant clone or SNP in that population, which may differ across populations.

SNP associations are replicated across populations. Replication of association signals is the sine qua non of association studies, and the fact that we observe this even between diverse populations and with small sample sizes highlights the relevance and robustness of the associations we detect. Gene expression is the basis for many crucial functions in the cell, so the relative contribution of these two types of variants is an indication of the nature of the mutational and natural selection processes that contribute to phenotypic diversity and divergence. It is, therefore, essential that we interrogate both SNPs and CNVs (of all types) to perform a comprehensive exploration of genetic effects on phenotypic variation and disease. It is possible that, if a larger number of SNPs were analyzed or a higher resolution of CNVs was available, we would observe more overlap between the effects attributed to CNVs and SNPs. However, the difficulty of designing robust SNP genotyping assays in structurally dynamic regions of the genome (26) suggests that even with more comprehensive interrogation of SNPs and CNVs, the overlap may not be high enough for one type of variation to be sufficient for exploring the genetic causes of disease. We have also demonstrated that it is not necessary to perform such studies with CNV calls or CNV genotypes, but it is possible to use filtered CGH log_2 ratios or any other type of high-quality quantitative signal that reflects underlying CNV. It has also become apparent that there are many more structural variants that contribute to phenotypic variation than our stringent criteria for what is a CNV reveal and that higher-resolution methods are necessary to elucidate their structure and function. Last, but not least, is the fact that we have only considered simple models of association in small samples, so it is very likely that if we apply more complex and realistic models (e.g., epistatic interactions) and or larger population samples, a larger number of effects would be revealed. The results presented here reinforce the idea that the complexity of functionally relevant genetic variation ranges from single nucleotides to megabases, and the full range of the effects of all of these variants will be best captured and interpreted by complete knowledge of the sequence of many human genomes. Until this is possible we need to survey all known types of genetic variation to maximize our understanding of human evolution, diversity, and disease.

References and Notes

1. B. E. Stranger *et al.*, *PLoS Genet.* **1**, e78 (2005).
2. V. G. Cheung *et al.*, *Nature* **437**, 1365 (2005).
3. S. Doss, E. E. Schadt, T. A. Drake, A. J. Lusis, *Genome Res.* **15**, 681 (2005).
4. R. B. Brem, L. Kruglyak, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 1572 (2005).
5. J. D. Storey, J. M. Akey, L. Kruglyak, *PLoS Biol.* **3**, e267 (2005).
6. M. F. Oleksiak, J. L. Roach, D. L. Crawford, *Nat. Genet.* **37**, 67 (2005).
7. S. A. Monks *et al.*, *Am. J. Hum. Genet.* **75**, 1094 (2004).
8. E. E. Schadt *et al.*, *Nature* **422**, 297 (2003).
9. E. J. Chesler *et al.*, *Nat. Genet.* **37**, 233 (2005).
10. L. Bystrykh *et al.*, *Nat. Genet.* **37**, 225 (2005).
11. E. T. Dermitzakis, B. E. Stranger, *Mamm. Genome* **17**, 503 (2006).
12. T. Pastinen, T. J. Hudson, *Science* **306**, 647 (2004).
13. International HapMap Consortium, *Nature* **437**, 1299 (2005).
14. L. Feuk, C. R. Marshall, R. F. Wintle, S. W. Scherer, *Hum. Mol. Genet.* **15** (suppl. 1), R57 (2006).
15. A. J. Iafrate *et al.*, *Nat. Genet.* **36**, 949 (2004).
16. J. Sebat *et al.*, *Science* **305**, 525 (2004).
17. E. Tuzun *et al.*, *Nat. Genet.* **37**, 727 (2005).
18. S. A. McCarroll *et al.*, *Nat. Genet.* **38**, 86 (2006).
19. D. F. Conrad, T. D. Andrews, N. P. Carter, M. E. Hurles, J. K. Pritchard, *Nat. Genet.* **38**, 75 (2006).
20. P. Stankiewicz, in *Genomic Disorders: The Genomic Basis of Disease*, J. R. Lupski, P. Stankiewicz, Eds. (Humana Press, Totowa, NJ, 2006), pp. 357–369.
21. D. A. Kleinjan, V. van Heyningen, *Am. J. Hum. Genet.* **76**, 8 (2005).
22. M. J. Somerville *et al.*, *N. Engl. J. Med.* **353**, 1694 (2005).
23. J. A. Lee *et al.*, *Ann. Neurol.* **59**, 398 (2006).
24. D. P. Locke *et al.*, *Am. J. Hum. Genet.* **79**, 275 (2006).
25. D. A. Hinds, A. P. Kloek, M. Jen, X. Chen, K. A. Frazer, *Nat. Genet.* **38**, 82 (2006).
26. R. Redon *et al.*, *Nature* **444**, 444 (2006).
27. GENEVAR—GENe Expression VARiation, www.sanger.ac.uk/genevar.
28. International HapMap Project, www.hapmap.org, release 16c.1.
29. The Copy Number Variation (CNV) Project Data Index, www.sanger.ac.uk/humgen/cnv/data.
30. R. W. Doerge, G. A. Churchill, *Genetics* **142**, 285 (1996).
31. G. Merla *et al.*, *Am. J. Hum. Genet.* **79**, 332 (2006).

32. We thank A. Clark and J. Pritchard for comments on earlier versions of the manuscript; M. Smith for assistance with software development; and M. Gibbs, J. Orwick, and C. Geringer for technical support. Funding was provided by the Wellcome Trust to E.T.D., M.E.H., P.D., C.T.S., and N.C., NIH to E.T.D. and S.T., Cancer Research U.K. to S.T. and N.T., the Leukemia and Lymphoma Society and the Brigham and Women's Hospital Department of Pathology to C.L., and the U.K. Medical Research Council (MRC) to M.D. S.T. is a Royal Society Wolfson Research Merit Award holder. S.W.S. is supported by grants from Genome Canada/Ontario Genomics Institute and is a Scholar of the Canadian Institutes of Health Research and the Howard Hughes Medical Institute.

Supporting Online Material
www.sciencemag.org/cgi/content/full/315/5813/848/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S5
References and Notes

24 October 2006; accepted 5 January 2007
10.1126/science.1136678

Evidence That Focal Adhesion Complexes Power Bacterial Gliding Motility

Tâm Mignot,[1*] Joshua W. Shaevitz,[2] Patricia L. Hartzell,[3] David R. Zusman[1*]

The bacterium *Myxococcus xanthus* has two motility systems: S motility, which is powered by type IV pilus retraction, and A motility, which is powered by unknown mechanism(s). We found that A motility involved transient adhesion complexes that remained at fixed positions relative to the substratum as cells moved forward. Complexes assembled at leading cell poles and dispersed at the rear of the cells. When cells reversed direction, the A-motility clusters relocalized to the new leading poles together with S-motility proteins. The Frz chemosensory system coordinated the two motility systems. The dynamics of protein cluster localization suggest that intracellular motors and force transmission by dynamic focal adhesions can power bacterial motility.

During the exhibition of gliding motility, bacteria move across solid surfaces without the use of flagella (1). Gliding motility is important for biofilm formation and bacterial virulence. Motility in *Myxococcus xanthus*, a Gram-negative rod-shaped bacterium, relies on two separate but coordinated motility engines. S motility is powered by type IV pili that are assembled at the leading cell pole; movement is produced as the pili bind to surface exopolysaccharides and are retracted, thereby pulling the cell forward (2). A motility, on the other hand, is not associated with pili or other obvious structures and is not well understood.

To investigate the A-motility system, we studied AglZ, a protein that is essential for A motility but dispensable for S motility (fig. S1, A and B) (3). AglZ is similar to FrzS, an S-motility protein that oscillates from one cell pole to the other when cells reverse direction (4) (fig. S1A). To track the localization of AglZ in moving cells, we constructed an *M. xanthus* strain containing a chimeric *aglZ-yfp* gene in place of the endogenous *aglZ* gene (fig. S2A). This chimeric gene encodes an AglZ yellow fluorescent protein (YFP) fusion protein that was stable and functional (fig. S2, B and C). We followed AglZ-YFP localization using time-lapse video microscopy. In fully mobile cells, AglZ-YFP was localized in

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA. [2]Department of Integrative Biology, University of California, Berkeley, CA 94720, USA. [3]Department of Microbiology, Molecular Biology, and Biochemistry, University of Idaho, Moscow, ID 83844, USA.

*To whom correspondence should be addressed. E-mail: tmignot@berkeley.edu (T.M.); zusman@berkeley.edu (D.R.Z.)

ordered clusters spanning the cell length; in stalled cells, it was localized at the leading cell pole (fig. S3).

To study the link between the localization of AglZ-YFP and motility, we focused our observations on AglZ-YFP in fully motile cells. These cells showed an ordered array of AglZ-YFP clusters spanning the cell body (Fig. 1A). As cells moved forward, AglZ-YFP clusters maintained fixed positions with respect to the agar substrate rather than to their relative positions in the cell (Fig. 1A). We analyzed the clusters by taking line scans of the fluorescence intensity along motility paths for successive movie frames (Fig. 1B). To identify the position of peaks in multiple frames, we calculated a thresholded line-scan average (Fig. 1B) (5). This analysis not only located the positions of clusters shown in Fig. 1A, but also had the sensitivity to find other common peaks that were difficult to identify when viewing images with the eye. In every cell that was examined (n = 30), the AglZ-YFP clusters remained fixed relative to the substratum. The only AglZ-YFP clusters that moved relative to the cell body were located at the leading pole, which suggests that new sites were assembled at that pole. The number of clusters per cell was dependent on cell length; clusters were disassembled when they reached the rear half of the cell (fig. S4).

We calculated the autocorrelation function of the line-scan averages from six different moving cells (Fig. 1C). Autocorrelations are useful for finding repeating patterns in a signal, such as the determination of the presence of periodicities buried under noise. The average of all six auto-

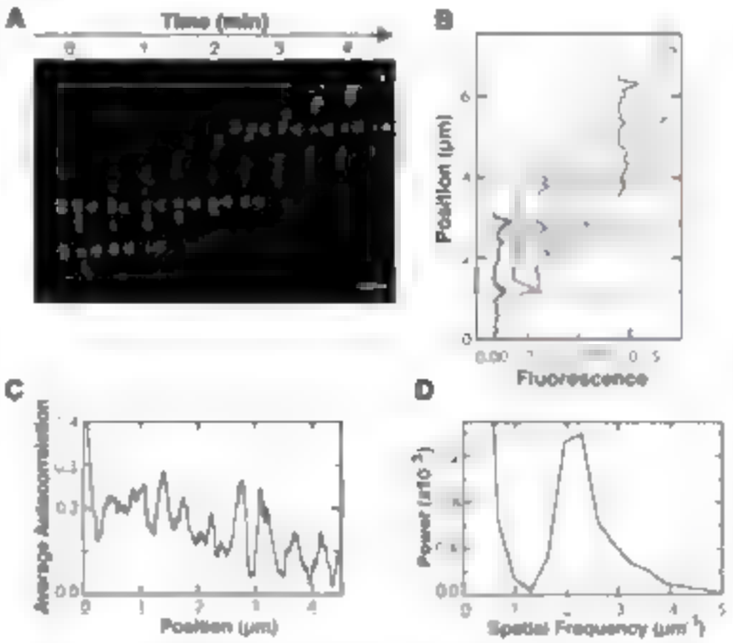


Fig. 1. AglZ-YFP localizes to periodic sites that remain fixed relative to the substratum. **(A)** AglZ-YFP localization in a cell moving at constant velocity. Fluorescence micrographs captured every 30 s are shown. Numbered arrowheads highlight selected bright fluorescence clusters. Scale bar, 1 μm. **(B)** Line scans of fluorescence intensity as a function of position are shown for each movie frame in (A). For display purposes, individual scans have been shifted horizontally with time. Gray bars represent the three highest peaks in the average line scan, matching the positions of clusters 1 to 3 in (A). Pink bars denote all additional peaks found in the average scan. **(C)** Quantitative analysis of the AglZ-YFP fluorescence distribution in moving cells. The autocorrelation function of the thresholded line scan from six moving cells was averaged and displayed. **(D)** Power spectral density of the autocorrelation function.

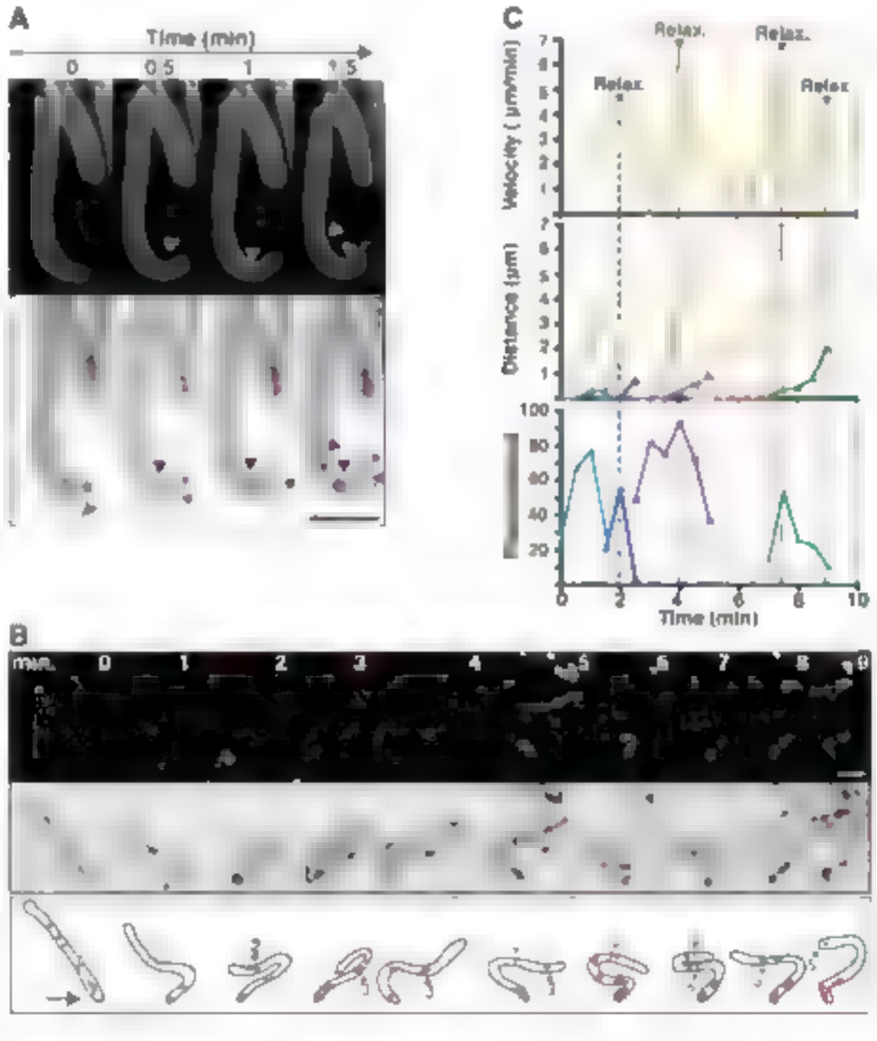


Fig. 2. AglZ-YFP localizes to transient adhesion sites. **(A)** AglZ-YFP fluorescence clusters in a cell that bends while in motion. (Top) Cells stained with the membrane dye FM4-64 are shown. (Bottom) An overlay of the membrane signal (gray) and the AglZ-YFP signal (magenta), which is artificially colored for better clarity, are shown. White and black arrowheads point to regions of cell-body curvature and localization of the YFP signal, respectively. Arrows indicate the direction of movement. Scale bar, 1 μm. **(B)** AglZ-YFP fluorescence clusters in a cell undergoing flailing motion. Fluorescence and overlaid phase micrographs (top and middle rows, respectively) are shown. Time intervals, 1 min. A cartoon representation (bottom row) shows the clusters numbered and color-coded for the analysis shown in (C). The arrow indicates the stuck leading pole. Scale bar, 2 μm. **(C)** Dynamic behavior of the AglZ-YFP fluorescence clusters in the cell shown in (B). Time intervals, 30 s. (Top) The velocity of the lagging pole over time is shown. Dotted lines mark the times where relaxation of the terminal bend (Relax.) is observed. The leading pole remained immobilized for the entire duration of the time lapse. (Middle) The distance traveled by the AglZ-YFP clusters, color-coded and numbered as in (B), over time is shown. 1, blue triangles; 2, blue diamonds; 3, purple squares; 4, pink squares; 5, green triangles. For each cluster, the distance traveled by the lagging pole (orange diamonds) during the same time interval was plotted to show that the clusters remain mostly fixed relative to the substratum. (Bottom) The relative fluorescence intensity of each cluster over time. The same color code as that used in the middle panel applies.

correlation functions displayed a clear periodicity (Fig. 1C), represented by a single large peak at a spatial frequency of 2.15 ± 0.03 μm^{-1} in its power spectral density (Fig. 1D). This frequency corresponded to a spatial periodicity of 466 ± 7 nm, which is very similar to the helical pitch of bacterial actin MreB as seen in both *Bacillus subtilis* [470 nm (6)] and *Escherichia coli* [460 ± 80 nm (7)].

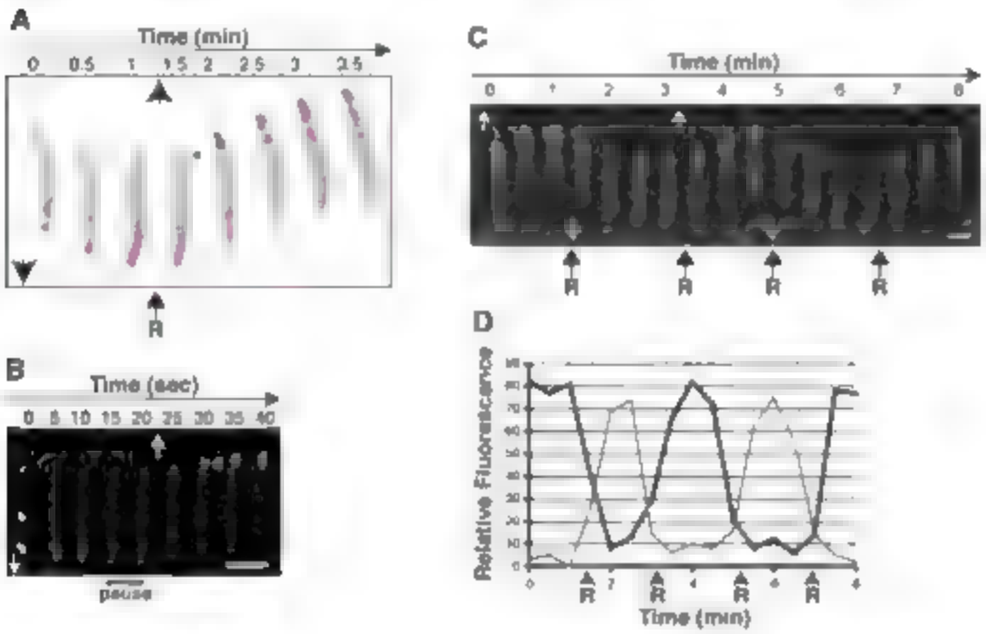


Fig. 3. AglZ-YFP oscillates from pole to pole upon cellular reversals. (**A**) AglZ-YFP localized to the new leading pole upon cellular reversals. Fluorescent micrographs of AglZ-YFP (magenta) and a representative reversing cell stained with FM4-64 (gray) were overlaid to show AglZ-YFP dynamics every 30 s. The black arrows inside the panel indicate the direction of movement. The arrowheads in (A) and (B) show the relocalization of AglZ-YFP at the new leading pole. R, reversal. (**B**) AglZ-YFP dynamics at the time of reversal. Fluorescence micrographs of a reversing cell captured every 5 s are shown. The 10-s delay is indicated by "pause." Scale bar, 2 μm. (**C**) AglZ-YFP oscillations in hyper-reversing cells. Fluorescent micrographs of a *frzCD*$^{-}$ cell that expresses AglZ-YFP captured every 30 s are shown. The white arrows in (B) and (C) indicate the direction of movement. Scale bar, 2 μm. (**D**) Quantitative fluorescence analysis of the cell presented in (C). The relative fluorescence intensities of each cell pole were measured in arbitrary units and plotted over time. The black line indicates the initial leading pole and the gray line indicates the initial trailing pole.

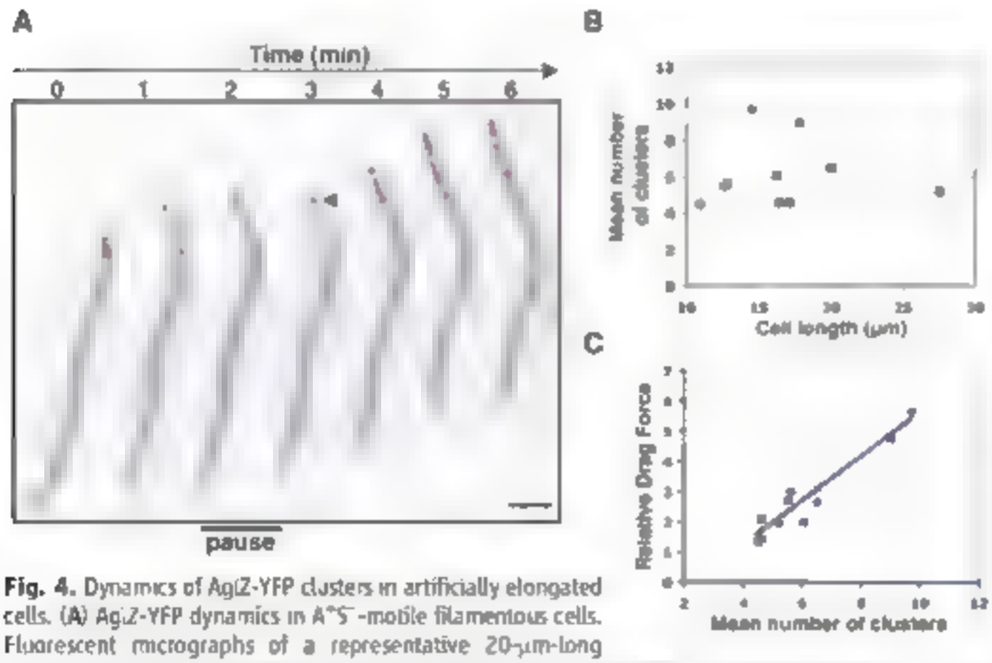


Fig. 4. Dynamics of AglZ-YFP clusters in artificially elongated cells. (**A**) AglZ-YFP dynamics in A^{+}S^{-}-motile filamentous cells. Fluorescent micrographs of a representative 20-μm-long cephalexin-treated cell stained with FM4-64 (gray) expressing AglZ-YFP (magenta). AglZ-YFP is only found distributed over the front part of the cell when the cell is in motion. The arrowhead indicates polar condensation of AglZ-YFP. "Pause" indicates times when the cell motion is stopped. Scale bar, 2 μm. (**B**) Relationship between cluster number and filamentous cell length. (**C**) Relationship between relative drag force overcome and cluster number in filamentous cells.

Because AglZ-YFP clusters remained fixed relative to the substratum as cells moved forward, they must be moving in the opposite direction to that of the cell. These dynamics suggest the formation of cytoskeleton-anchored transient surface adhesions that could power cellular movement. We analyzed cells that bend while in motion to investigate whether AglZ-YFP clusters could indicate sites of transient adhesion between the cell surface and the substratum. First, in a moving cell that was progressively bending, two distinct sites of curvature were observed (Fig. 2A). If AglZ-YFP localizes at sites that mediate adhesion with the substratum, then clusters should accumulate at sites of maximum bending. As expected, accumulation of AglZ-YFP was specifically observed at each site. Second, cells sometimes became stuck at their leading end, which prevented forward movement; these cells bent into C or S shapes as the active A engine pushed against the flexible cell wall. It has been proposed that these flailing motions are due to a motor pushing from the lagging end of the cell (8). However, if multiple motor-coupled adhesion complexes pushed cells forward, flailing would also occur. In that situation, cell bends would form between adhesion complexes. In a flailing cell expressing AglZ-YFP (Fig. 2B), as the front of the cell became fixed, the still-motile cell adopted a right-handed C shape and then relaxed to adopt a left-handed U shape that transitioned into an S shape to become a right-handed U shape again (neighboring cells probably block full relaxation between 2 to 3 min and 8 to 9 min). Cell shape correlated with the pattern of AglZ-YFP clusters. At each bend, enlarged AglZ-YFP fluorescence clusters were not observed at the maximum inflection points, but rather between the bends. The fluorescent clusters maintained relatively fixed positions with respect to the underlying substratum as the cell body appeared to move through them (Fig. 2, B and C). Furthermore, as the AglZ-YFP clusters faded and dispersed at the lagging cell pole, when cell bends relaxed, a spike in velocity was observed, which suggests that dispersal of the AglZ-YFP clusters indeed removed adhesion constraints (Fig. 2C). Thus, the cell body appears to move through "focal adhesion" sites where AglZ-YFP accumulates.

The dynamics of AglZ-YFP localization (Fig. 1A) suggest that A-engine clusters are assembled at the leading cell pole and disassembled at the lagging end. It follows then that, during cellular reversals, proteins of the A-motility system should be shifted to the new leading pole along with S-motility components (4). Indeed, upon cellular reversal, AglZ-YFP localized rapidly to the new leading pole (Fig. 3A and table S2). This was a two-step process. Immediately before reversal, the cell paused for 10 s, and AglZ-YFP became diffuse (Fig. 3B). The cell then reversed, and AglZ-YFP

localized to the new leading pole after a 10-s delay. Thus, AglZ may not trigger polar switching, but rather might be assembled at the new leading pole along with other motility components.

The Frz chemosensory pathway, which regulates cell reversals, controls FrzS translocation to the leading pole when cells reverse direction (4). In a nonreversing *frzE* mutant, AglZ-YFP localization never switched poles. In contrast, a hyperreversing *frzCD* mutant showed very frequent reversals that were always followed by AglZ-YFP polar switching (Fig. 3, C and D, and table S2). In this *frzCD* strain, AglZ-YFP was mostly polar in distribution, and ordered intracellular fluorescence clusters were only observed transiently. Thus, cellular reversals result from the concerted switching of both A- and S-motility components to the new leading pole (4). Coordination of these two engines must be achieved through the signaling activity of a common pathway.

Although AglZ-YFP localized to transient adhesion sites, it is unclear whether the force that produces locomotion is generated at those sites. To address this question, we investigated the motility of cells treated with the antibiotic cephalexin. Cephalexin-treated cells, which elongate up to 10 times their natural length, showed almost normal A motility but greatly reduced S motility, which suggests that the A engine is distributed along the cell body whereas the S engine is polar (9). We also observed that 10- to 30-μm-long A^+S^--motile cephalexin-treated cells moved with velocities that were independent of cell length. The localization of AglZ-YFP was also correlated with the activity of the A engine in these cells (Fig. 4A). In these moving filaments, AglZ-YFP was localized in clusters that were distributed in the front part of the cell, whereas the back of the cell was largely depleted of clusters (Fig. 4A); consequently, the number of clusters per cell was largely independent of cell length (Fig. 4B). Thus, in the filaments, we could test whether force was produced at the sites where AglZ-YFP accumulates by analyzing the relationship between the number of sites and the "drag force overcome" [i.e., the force necessary to power the motility of a cell of given cell length and velocity (5)]. Indeed, the drag force overcome was proportional to the number of clusters in filamentous cells (Fig. 4C), indicating that motility force seems to be produced at the adhesion sites; these characteristics are similar to eukaryotic focal adhesions where both adhesion and force are generated (10).

Previously, a "slime gun" model for gliding motility was proposed because, in several bacterial species, motility is correlated with the secretion of slime through pores (nozzles) located in the outer membrane (11, 12), and a biophysical model suggested that the hydration of slime within the nozzles could generate sufficient force to propel bacteria forward (11). Our results are consistent with an alternate model, whereby intracellular motor complexes that connect to both membrane-spanning adhesion complexes and to the cytoskeleton power motility by pushing against the substratum and moving the cell body forward, much like focal adhesion–based traction of apicomplexan gliding motility in eukaryotic organisms (fig. S5) (10, 13). The periodicity of the AglZ-YFP clusters strongly implies the existence of a continuous helical filament that spans the length of the cell. Thus, the action of the motor complexes may induce the cell body to rotate as it pulls the cell forward (fig. S5). Rotation of the cell body as cells move has been shown for *Cytophaga* sp., an organism that also secretes slime during motility (14). Slime secretion may be part of the motility system, by supplying additional power for movement and/or adhesion or lubrication of the interface between the cell body and the substratum (10). The dynamics of AglZ-YFP and FrzS green fluorescent protein (GFP) fusion protein (4) suggest how the A and S engines might be coordinated. A and S complexes oscillate so that they are targeted together to the new leading pole upon cellular reversal, which is synchronized by the Frz chemosensory system.

References and Notes

1. A. M. Spormann, *Microbiol. Mol. Biol. Rev.* **63**, 621 (1999).
2. Y. Li *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5443 (2003).
3. R. Yang *et al.*, *J. Bacteriol.* **186**, 6168 (2004).
4. T. Mignot, J. P. Merlie, D. R. Zusman, *Science* **310**, 855 (2005).
5. Materials and methods are available as supporting material on *Science* Online.
6. H. J. Defeu Soufo, P. L. Graumann, *BMC Cell Biol.* **6**, 10 (2005).
7. T. Kruse, J. Moller-Jensen, A. Lobner-Olesen, K. Gerdes, *EMBO J.* **22**, 5283 (2003).
8. C. W. Wolgemuth, *Biophys. J.* **89**, 945 (2005).
9. H. Sun, Z. Yang, W. Shi, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 15178 (1999).
10. M. A. Wozniak, K. Modzelewska, L. Kwong, P. J. Keely, *Biochim. Biophys. Acta* **1692**, 103 (2004).
11. C. Wolgemuth, E. Hoiczyk, D. Kaiser, G. Oster, *Curr. Biol.* **12**, 369 (2002).
12. E. Hoiczyk, W. Baumeister, *Curr. Biol.* **8**, 1161 (1998).
13. J. Baum, A. T. Papenfuss, B. Baum, T. P. Speed, A. F. Cowman, *Nat. Rev. Microbiol.* **4**, 621 (2006).
14. S. L. Godwin, M. Fletcher, R. P. Burchard, *J. Bacteriol.* **171**, 4589 (1989).

15. We thank Y. Inclan for the construction of *pXe*::*aglZ* and comments on the manuscript, and J. Merlie, R. Losick, and G. Oster for helpful discussions and comments. This research was supported by a grant from NIH to D.R.Z. (GM20509).

Supporting Online Material
www.sciencemag.org/cgi/content/full/315/5813/853/DC1
Materials and Methods
SOM Text
Figs. S1 to S5
Tables S1 and S2
References

7 November 2006; accepted 12 December 2006
10.1126/science.1137223

Apoptosis Initiated When BH3 Ligands Engage Multiple Bcl-2 Homologs, Not Bax or Bak

Simon N. Willis,[1] Jamie I. Fletcher,[1] Thomas Kaufmann,[1] Mark F. van Delft,[1,2] Lin Chen,[1] Peter E. Czabotar,[1] Helen Ierino,[1] Erinna F. Lee,[1,2] W. Douglas Fairlie,[1] Philippe Bouillet,[1] Andreas Strasser,[1] Ruth M. Kluck,[1] Jerry M. Adams,[1]* David C. S. Huang[1]*†

A central issue in the regulation of apoptosis by the Bcl-2 family is whether its BH3-only members initiate apoptosis by directly binding to the essential cell death mediators Bax and Bak, or whether they can act indirectly, by engaging their pro-survival Bcl-2–like relatives. Contrary to the direct activation model, we show that Bax and Bak can mediate apoptosis without discernable association with the putative BH3-only activators (Bim, Bid, and Puma), even in cells with no Bim or Bid and reduced Puma. Our results indicate that BH3-only proteins induce apoptosis at least primarily by engaging the multiple pro-survival relatives guarding Bax and Bak.

Whether a cell dies in response to diverse developmental cues or cellular stresses is determined largely by interactions between three factions of the Bcl-2 protein family (1). Two factions promote apoptosis: The BH3-only proteins (including Bim, Bid, Puma, Bad, and Noxa) sense cellular damage, and Bax and Bak are critical downstream mediators of apoptosis because their combined absence abolishes most apoptotic responses (2, 3). When activated, Bax and Bak permeabilize the outer mitochondrial membrane, freeing proapoptogenic factors such as cytochrome c, which promote activation of the proteases (caspases) that mediate cellular demolition. Activation of Bax and Bak is opposed by the pro-survival faction: Bcl-2, $Bcl-x_L$, Bcl-w, Mcl-1, and A1. These guardians are inactivated when BH3-only proteins insert their BH3 domain into a groove on the pro-survival proteins (1).

A widely embraced model for the initiation of apoptosis (4) (Fig. 1A) proposes that a subset of BH3-only proteins termed "activators"—namely

[1]The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3050, Australia. [2]Department of Medical Biology, University of Melbourne, Parkville, Victoria 3010, Australia.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: huang_d@wehi.edu.au

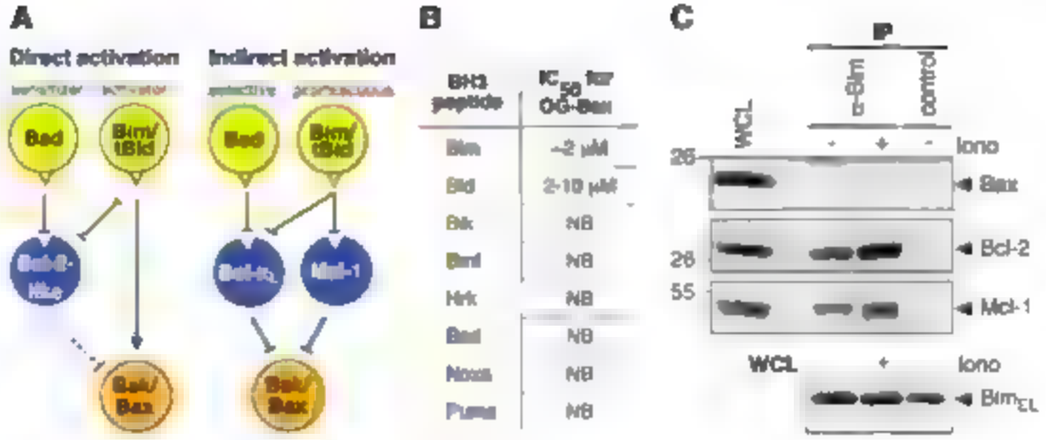


Fig. 1. BH3 binding to Bax and Bak. **(A)** Models for activation of Bax or Bak. **(B)** Relative affinities of BH3 peptides for octyl glucoside (OG)–treated Bax, determined in solution competition assays (*8*, *10*). IC_{50}, median inhibitory concentration. NB, no binding at 10 μM. **(C)** Coimmunoprecipitation tests in Triton X-100–containing lysates of healthy thymocytes (–) or 24 hours after treatment with ionomycin (+), a Bim-dependent stimulus (*14*). Bim-containing complexes, isolated using an antibody (3C5) that recognizes all Bim isoforms, were examined for Bax, Bcl-2, or Mcl-1 (upper three panels). Control, isotype-matched antibody immunoprecipitation; WCL, whole-cell lysates; IP, immunoprecipitation.

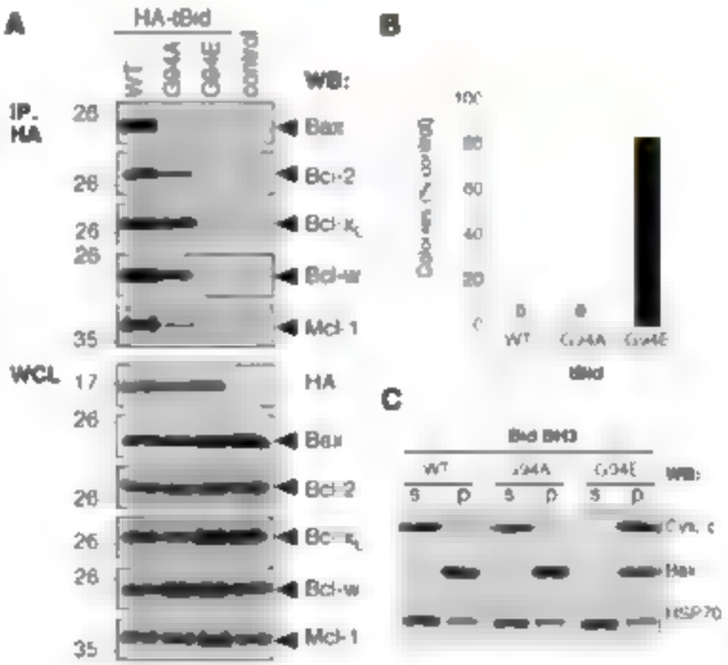


Fig. 2. Binding to Bax is not required for killing by tBid. **(A)** In Triton X-100–containing 293T lysates, the binding of overexpressed hemagglutinin (HA)–tagged tBid [wild type (WT) or its G94 mutants to endogenous Bax, or of overexpressed pro-survival proteins, was tested by coimmunoprecipitation with an antibody to HA and immunoblotting with the indicated antibodies. Lower panels, protein expression in WCLs. Control, immunoprecipitation from untransfected cells. **(B)** Colony formation 6 days after plating of equivalent numbers of retrovirally infected MEFs (*10*), normalized to the number of colonies formed with the parental retrovirus. Data represent means ± SD from three independent experiments. **(C)** Cytochrome *c* release elicited by BH3 peptides. Permeabilized MEFs incubated with each BH3 peptide (10 μM) were separated into soluble (S) and sedimented (P, pellet) fractions that were probed with antibodies to cytochrome *c*, Bax (membrane marker), or HSP70 (cytosolic marker).

Bim and Bid (*4*–*6*) and, perhaps, Puma (*7*) directly engage Bax or Bak. The other BH3-only proteins, termed "sensitizers" (for example, Bad and Noxa), purportedly act only by displacing the activators from the pro-survival proteins, allowing the activators to bind Bax and Bak. An alternative indirect model (*8*, *9*) (Fig. 1A), supported here, posits that BH3-only proteins activate Bax or Bak not by binding either one but indirectly by engaging the multiple pro-survival relatives that constrain them. In this scenario, Bim, Bid, and Puma are highly potent because they engage all pro-survival proteins, whereas the others are less so because they bind only selective subsets (*8*).

We first investigated whether BH3 peptides associate with Bax in vitro (*10*). The Bim and Bid BH3 peptides failed to bind native Bax (fig. S1A), and even a large excess of the Bim peptide did not promote Bax oligomerization (fig. S2). In non-ionic detergents that alter the Bax conformation (fig. S3) (*11*), perhaps mimicking an activation step, they did bind (fig. S1A). However, their affinities for Bax (in the micromolar range) (Fig. 1B) appeared to be lower than those (in the nanomolar range) for the pro-survival proteins (*8*) (fig. S1C). Thus, Bim and Bid peptides bound weakly and only after Bax had changed conformation. No other BH3 peptide, including that of Puma, bound even to activated Bax (Fig. 1B).

Coimmunoprecipitation studies extended these observations to full-length BH3-only proteins in cell lysates (fig. S4) (*10*). In lysates of human embryonic kidney 293T cells prepared with the detergent Triton X-100, BH3-only proteins bound to Bcl-2 or other pro-survival proteins (*8*) but not to endogenous Bak (fig. S4). Similarly, in lysates made with CHAPS, a detergent that maintains Bax as a soluble monomer (*11*), no BH3-only protein bound Bax. In Triton X-100 lysates, Bax did bind detectably to the truncated Bid (tBid) produced by caspases, but not to full-length Bid (fig. S4A), which has its BH3 domain buried (*12*). Bim_{EL} and Bim_L, the only Bim isoforms detectable in multiple cell types (fig. S5), failed to bind Bax (fig. S4B) (*13*), though both killed cells efficiently (*8*). Bax did bind the very minor Bim_S isoform (fig. S4B) and some unusual Bim variants confined to certain primates (fig. S6). However, no Bax was detectable in endogenous Bim protein complexes (fig. 1C), even those from mouse thymocytes treated with ionomycin, which kills via a Bim-dependent pathway (*14*). In contrast, the binding of Bim to endogenous pro-survival proteins was evident, and this binding increased after cellular insults (Fig. 1C) (*15*, *16*). Thus, Bim did not form physiologically relevant complexes with Bax or Bak.

To test whether cell killing by tBid or Bim_S requires their binding to Bax, we used a mutation, glycine to alanine [$Gly^{94} \rightarrow Ala^{94}$ (G94A)], within their BH3 domains that abrogated binding to Bax but not to the pro-survival proteins (Fig. 2 and figs. S7 to S9). Despite its inability to bind Bax, tBid G94A killed mouse embryo fibroblasts (MEFs) and mouse myeloid cells as effectively as the wild-type protein, in both short-term viability (figs. S7 and S8) and clonogenic (Fig. 2B) assays. Our studies with MEFs lacking Bax or Bak (fig. S9) showed that either Bak or Bax could mediate the killing. Furthermore, the Bid G94A BH3 peptide, like the wild-type peptide, promoted cytochrome *c* release from mitochondria in vitro (Fig. 2C) (*10*). The equivalent Bim_S mutant gave similar results (figs. S8 and S9).

Thus, tBid and Bim can induce apoptosis without binding Bax or Bak. We inferred that these proteins and their G-to-A mutants caused cell death by binding to and neutralizing diverse

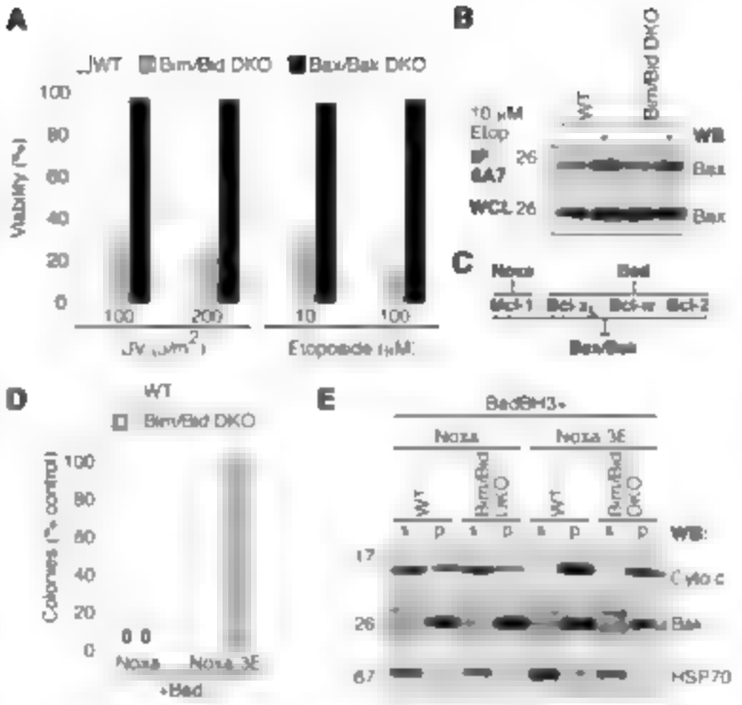


Fig. 3. Apoptosis in the absence of both Bim and Bid. (**A**) Viability of cells lacking Bim and Bid after UV irradiation or etoposide treatment. (**B**) Bax activation without Bim-Bid. Immunoprecipitation with monoclonal antibody 6A7, which recognizes only activated Bax (*11*), from WT or *bim*$^{-/-}$*bid*$^{-/-}$ MEFs treated for 24 hours with etoposide. (**C** and **D**) Killing by sensitizer BH3-only proteins. (**C**) BH3 specificity. (**D**) Colony formation after Bad retroviral infection of MEFs stably expressing Noxa (to target Mcl-1) or the inert Noxa mutant 3E (*9*). (**E**) Cytochrome *c* release without Bim or Bid. Digitonin-permeabilized MEFs expressing Noxa or Noxa 3E were incubated with 10 μM Bad BH3 peptide before separation into soluble (s) and pellet (p) fractions. Data in (A) and (D) represent means ± SD from three independent experiments.

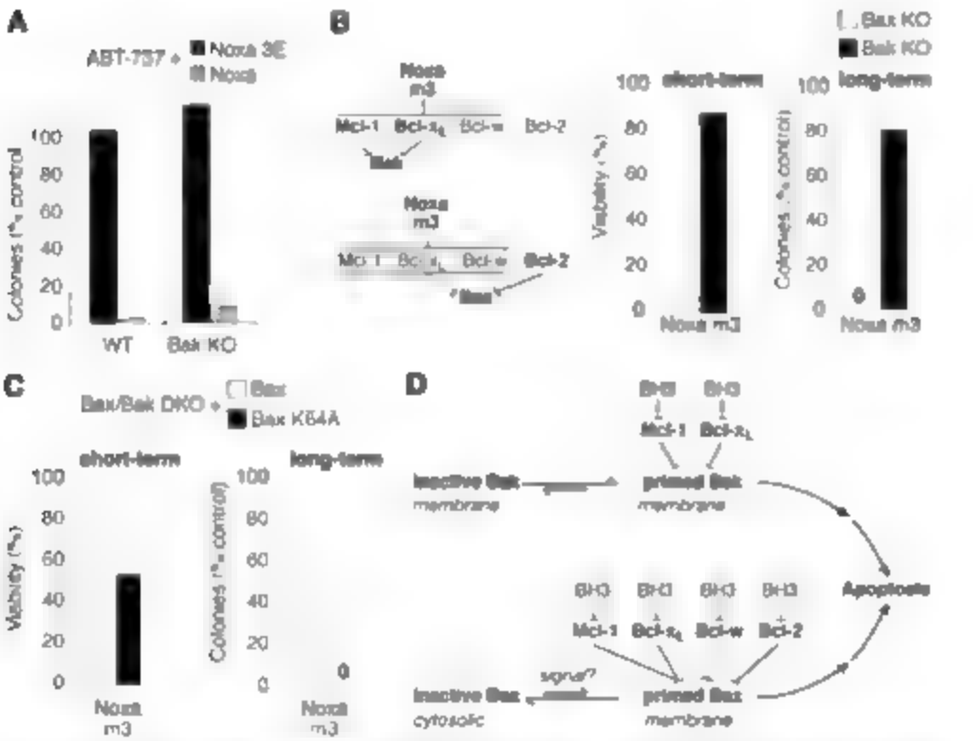


Fig. 4. Control of Bax by multiple pro-survival relatives. (**A**) Colony formation by WT or Bax-expressing (*bak*$^{-/-}$) MEFs stably expressing Noxa or the inert Noxa 3E (*9*) and treated with ABT-737 (Bad BH3 mimetic) to inactivate diverse pro-survival proteins. (**B**) Short-term viability or colony formation for *bax*$^{-/-}$ or *bak*$^{-/-}$ MEFs infected with Noxa m3. The model depicts Bak (*9*) or Bax activation upon neutralization of pro-survival proteins by Noxa m3. (**C**) Short term or clonogenic assays after Noxa m3 infection of *bax*$^{-/-}$ *bak*$^{-/-}$ MEFs reconstituted with Bax or the K64A mutant. (**D**) Regulation of Bak or Bax by multiple pro-survival relatives. Until inhibited by BH3 ligands, the indicated pro-survival proteins can sequester primed Bak or Bax (possibly a form with its BH3 exposed) on intracellular membranes. Data in (A) to (C) represent means ± SD from three independent experiments.

pro-survival proteins (*8*). Indeed, the corresponding glycine-to-glutamate mutants, which do not bind most (or any) pro-survival family members, did not kill (Fig. 2 and figs. S7 to S9).

If Bim and Bid [the major candidate activator BH3-only proteins (*4*, *6*)] were essential for apoptosis, as the direct-activation model implies, mice lacking both proteins (*10*) should exhibit developmental abnormalities, like Bax-Bak deficient animals, which usually die as neonates (*2*). Instead, mice lacking both Bim and Bid were born at the expected Mendelian frequency (fig. S10), appeared normal (e.g., no webbing between the toes), and remained healthy for over 12 months. Moreover, comparison of the sensitivity to DNA damage of immortalized MEFs from Bim-Bid double-knockout (DKO), wild-type, or Bax-Bak DKO embryos (fig. S11) revealed that, whereas Bax-Bak deficiency conferred resistance to ultraviolet (UV) irradiation and etoposide, the absence of both Bim and Bid neither abrogated apoptosis nor prevented Bax activation (Fig. 3, A and B).

Because the direct-activation model (Fig. 1A) proposes that only an activator BH3-only protein can kill cells, we tested whether the sensitizers Bad and Noxa (*4*) could kill cells lacking both Bim and Bid. Although overexpression of Bad and Noxa together killed wild-type MEFs (*8*) and matched Bax and Bak (Fig. 1B and fig. S8), we could not exclude the possibility that they killed by displacing endogenous Bim or Bid. However, even in the absence of both Bim and Bid, the Noxa-Bad combination (Fig. 3, C to E), which neutralizes all four pro-survival proteins expressed in MEFs (Bcl-2, Bcl-x_L, Bcl-w, and Mcl-1), still induced apoptosis and cytochrome *c* release in vitro, as did the Bad-like BH3 mimetic ABT-737 (*17*) plus Noxa (fig. S12). Mouse liver mitochondria yielded similar results (fig. S13).

Although one study reported that Puma bound Bax (*7*), its putative activator role is disputed (*5*, *6*). We and others (*18*) have detected no association of Puma with Bax or Bak (Fig. 1B and fig. S8C). Nevertheless, to assess whether Puma has any role as an activator, we decreased Puma expression with RNA interference in *bim*$^{-/-}$ *bid*$^{-/-}$ MEFs (fig. S14). As expected (*19*), its reduction impaired etoposide-induced death. Nevertheless, in cells retaining only a small amount of Puma and no Bim or Bid, the Noxa-ABT-737 combination still efficiently provoked apoptosis and cytochrome *c* release (fig. S14).

We conclude that BH3-only proteins can initiate apoptosis without binding Bax or Bak, although the cooperativity with Bax of tBid (or of a stapled Bad BH3 peptide in lysing protein-free liposomes (*5*, *20*, *21*) may indicate that direct activation contributes under some circumstances. We (*8*, *9*) and others (*22*) have suggested that all BH3-only proteins function by engaging their pro-survival relatives and overcoming their block to Bax or Bak activation (Fig. 1A). Mcl-1 and Bcl-x_L (but not Bcl-2) can constrain Bak by sequestering it, because enforced coexpression of Bad (to target Bcl-x_L) and Noxa (to target Mcl-1)

elicited Bak-dependent apoptosis without any additional stimulus (9).

Which of the multiple pro-survival proteins that can bind Bax (fig. S15A) can functionally restrain it? Mcl-1 must, because neutralizing Mcl-1 by enforced Noxa expression rendered MEFs containing only Bax (Bak KO cells) sensitive to the Bad BH3 mimetic ABT-737 (Fig. 4A), which inactivates Bcl-2, Bcl-x_L, and Bcl-w (17). To identify other pro-survival antagonists of Bax (10), we used Noxa mutant m3, which binds Mcl-1, Bcl-x_L, and Bcl-w, but not Bcl-2 (8) (Fig. 4B). It kills cells expressing only Bak (9), but not Bax-expressing ones (i.e., Bak KO cells) (Fig. 4B). Because Noxa m3 spares Bcl-2, Bcl-2 is implicated as another check on Bax. We tested this hypothesis using a functional Bax mutant [Lys[64] ⟶ Ala[64] (K64A)] (fig. S15, C and D) that loses Bcl-2 binding but still binds the other pro-survival proteins (fig. S15B). The ability of Noxa m3 to kill MEFs that express K64A but not wild-type Bax (Fig. 4C) confirms that Bcl-2 can prevent Bax activation. Hence, Bax is held in check by Mcl-1, Bcl-2, and either Bcl-x_L or Bcl-w, or by all four (Fig. 4D). Genetically, both Bcl-2 and Bcl-x_L can constrain Bax, because its concomitant loss precludes the lymphoid hypoplasia (23) or neuronal apoptosis (24) caused by their respective deficiencies. Thus, whereas only certain pro-survival proteins keep Bak in check (9), all of these proteins probably inhibit Bax (Fig. 4D). Presumably, they bind a "primed" conformer of Bak or Bax, a form that may normally be present at low levels or may be formed early in apoptotic signaling.

Our findings emphasize that the central role of the mammalian pro-survival Bcl-2 proteins, like their virally encoded orthologs (25), is to inhibit Bax and Bak (Fig. 4D) (8, 9). Thus, BH3-only proteins trigger apoptosis indirectly through Bax or Bak by neutralizing all of the relevant pro-survival proteins, allowing the activation of Bax and Bak to proceed.

References and Notes

1. J. M. Adams, *Genes Dev.* **17**, 2481 (2003).
2. F. Lindsten *et al.*, *Mol. Cell* **6**, 1389 (2000).
3. E. H. Cheng *et al.*, *Mol. Cell* **8**, 705 (2001).
4. A. Letai *et al.*, *Cancer Cell* **2**, 183 (2002).
5. T. Kuwana *et al.*, *Mol. Cell* **17**, 525 (2005).
6. M. Certo *et al.*, *Cancer Cell* **9**, 351 (2006).
7. P. F. Cartron *et al.*, *Mol. Cell* **16**, 807 (2004).
8. L. Chen *et al.*, *Mol. Cell* **17**, 393 (2005).
9. S. N. Willis *et al.*, *Genes Dev.* **19**, 1294 (2005).
10. Materials and methods are available as supporting material on *Science* Online.
11. Y.-T. Hsu, R. J. Youle, *J. Biol. Chem.* **273**, 10777 (1998).
12. J. J. Chou, H. Li, G. S. Salvesen, J. Yuan, G. Wagner, *Cell* **96**, 615 (1999).
13. M. Marani, T. Tenev, D. Hancock, J. Downward, N. R. Lemoine, *Mol. Cell. Biol.* **22**, 3577 (2002).
14. P. Bouillet *et al.*, *Science* **286**, 1735 (1999).
15. Y. Zhu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7681 (2004).
16. P. Bouillet *et al.*, *Nature* **415**, 922 (2002).
17. T. Oltersdorf *et al.*, *Nature* **435**, 677 (2005).
18. L. Wang, P. Wang, A. Bank, J. Yu, L. Zhang, *J. Biol. Chem.* **281**, 14034 (2006).
19. A. Villunger *et al.*, *Science* **302**, 1036 (2003).
20. T. Kuwana *et al.*, *Cell* **111**, 331 (2002).
21. L. D. Walensky *et al.*, *Mol. Cell* **24**, 199 (2006).
22. S. C. Ruffolo, G. C. Shore, *J. Biol. Chem.* **278**, 25039 (2003).
23. C. M. Knudson, S. J. Korsmeyer, *Nat. Genet.* **16**, 358 (1997).
24. K. S. Shindler, C. B. Latham, K. A. Roth, *J. Neurosci.* **17**, 3112 (1997).
25. A. Cuconati, K. Degenhardt, R. Sundararajan, A. Anschel, E. White, *J. Virol.* **76**, 4547 (2002).

26. The authors thank P. Colman, S. Cory, C. Day, G. Dewson, K. Newton, and D. Vaux for discussions. Abbott Laboratories (S. Fesik, S. Rosenberg, S. Elmore, and A. Shoemaker) for ABT 737; S. Coyne, M. Evangelista, B. Helbert, B. Kolevski, L. Parma, T. Pham, M. Robati, L. Tai, and L. Zhang for technical assistance; D. Cooper, M. Iannarella, and G. Siciliano for animal husbandry; and R. Anderson, G. Dewson, M. Hemann, S. Korsmeyer, S. Lowe, E. Michalak, H. Puthalakath, M. Suzuki, C. Thompson, and W. Welch for reagents. Our work is supported by scholarships, fellowships, and grants from the Cancer Council Victoria (S.N.W., E.F.L., and R.M.K.), the Swiss National Science Foundation (T.K.), the Roche Research Foundation (T.K.), the Novartis Foundation (T.K.), a Melbourne International Research Scholarship (M.F.v.D.), the Wellcome Trust (R.M.K.), the Sylvia & Charles Viertel Charitable Foundation (P.B. and D.C.S.H.), the Australian National Health and Medical Research Council (program grant 257502 and fellowships to W.D.F., A.S., J.M.A., and D.C.S.H.), the U.S. National Cancer Institute (grants CA80188 and CA43540), and the Leukemia & Lymphoma Society (grant SCOR 7015-02).

Supporting Online Material
www.sciencemag.org/cgi/content/full/315/5813/856/DC1
Materials and Methods
Figs. S1 to S15
References

1 August 2006; accepted 6 December 2006
10.1126/science.1133289

CPD Damage Recognition by Transcribing RNA Polymerase II

Florian Brueckner,[1–3] Ulrich Hennecke,[1,2] Thomas Carell,[1,2]* Patrick Cramer[1–3]*

Cells use transcription-coupled repair (TCR) to efficiently eliminate DNA lesions such as ultraviolet light–induced cyclobutane pyrimidine dimers (CPDs). Here we present the structure-based mechanism for the first step in eukaryotic TCR, CPD-induced stalling of RNA polymerase (Pol) II. A CPD in the transcribed strand slowly passes a translocation barrier and enters the polymerase active site. The CPD 5′-thymine then directs uridine misincorporation into messenger RNA, which blocks translocation. Artificial replacement of the uridine by adenosine enables CPD bypass; thus, Pol II stalling requires CPD-directed misincorporation. In the stalled complex, the lesion is inaccessible, and the polymerase conformation is unchanged. This is consistent with nonallosteric recruitment of repair factors and excision of a lesion-containing DNA fragment in the presence of Pol II.

Ultraviolet light damages cellular DNA by inducing dimerization of adjacent pyrimidines in a DNA strand. The resulting cyclobutane pyrimidine dimer (CPD) lesions can block transcription and replication and are a major cause of skin cancer (1). Cells eliminate CPDs by nucleotide excision repair (NER). A very efficient NER subpathway is transcription-coupled DNA repair (TCR), which specifically removes lesions from the DNA strand transcribed by RNA polymerase (Pol) II (2). Pol II stalls when a CPD in the DNA template strand reaches the enzyme active site (3, 4). Pol II stalling apparently triggers TCR by recruitment of a transcription-repair coupling factor [Rad26 in yeast, Cockayne syndrome protein B (CSB) in humans] and factors required for subsequent steps of NER, including TFIIH, which comprises helicases that unwind DNA (5). Endonucleases then incise the DNA strand on either side of the lesion, which results in a 24- to 34-nucleotide fragment (6–8). The obtained DNA gap is subsequently filled by DNA synthesis and ligation (9, 10).

To elucidate CPD recognition by transcribing Pol II, we carried out a structure-function analysis of elongation complexes containing in the template strand a thymine-thymine CPD. Elongation complexes were reconstituted from the 12-subunit *Saccharomyces cerevisiae* Pol II and nucleic acid scaffolds, as previously described (11), except that the mobile upstream DNA and the nontemplate strand in the transcription bubble were omitted (SOM text). A chemical analog of a CPD lesion was incorporated at register +2/+3 of the template strand, directly downstream of position +1, which denotes the substrate addition site (scaffold A, Fig. 1A and SOM text) (12). The crystal structure of the resulting elongation complex A was determined (Fig. 1B) (12), and the register of the nucleic acids was unambiguously defined by

[1]Munich Center for Integrated Protein Science CiPS[M], [2]Department of Chemistry and Biochemistry, [3]Gene Center Munich, Ludwig-Maximilians-Universität München, Feodor-Lynen-Strasse 25, 81377 Munich, Germany.

*To whom correspondence should be addressed. E-mail: thomas.carell@cup.uni-muenchen.de (T.C.); cramer@lmb.uni-muenchen.de (P.C.)

bromine labeling (table S1 and fig. S1C). The overall structure of complex A was nearly identical to the complete Pol II elongation complex (*11*) and very similar to elongation complex structures of the Pol II core enzyme (*13*, *14*) (SOM text). The template strand enters the active site, and continues into an 8 base pair hybrid duplex with RNA, which occupies the upstream positions −1 to −8 (Fig. 1B and SOM text). It contrasts to the earlier free elongation complex (*11*), downstream DNA entering the cleft is mobile, which indicated that the CPD at positions +2 +3 loosens the grip on downstream DNA (SOM text).

To investigate Pol II stalling at the CPD lesion, we incubated complex A with nucleoside triphosphate (NTP) substrates, followed RNA extension by fluorescence-monitored capillary electrophoresis, and identified the RNA products by mass spectrometry (Figs. 2 and 3) (*12*). After incubation with a physiological concentration of 1 mM NTPs for 1 hour, the RNA was extended by three nucleotides (Fig. 2A), but not any further. Thus, the complex stalled after nucleotide incorporation opposite both CPD thymines. A time course showed that the first incorporation event was fast, consistent with a [illegible] site in complex A (Figs. 1, A and B, and 2B) (SOM text). The second and third incorporation events, however, were progressively slower, with rate constants of ~16 per hour and 2.4 per hour, respectively (Fig. 2, A and B).

Structural considerations suggest that the second incorporation event is slower because translocation of the CPD from position +2 +3 to position +1 +2 is disfavored. Template bases in positions [illegible] and −1 are twisted against each other by 90° in the undamaged elongation complex (*11*), but twisting the CPD thymines is impossible because they are covalently linked. To test this, we introduced CPD at positions +1 +2

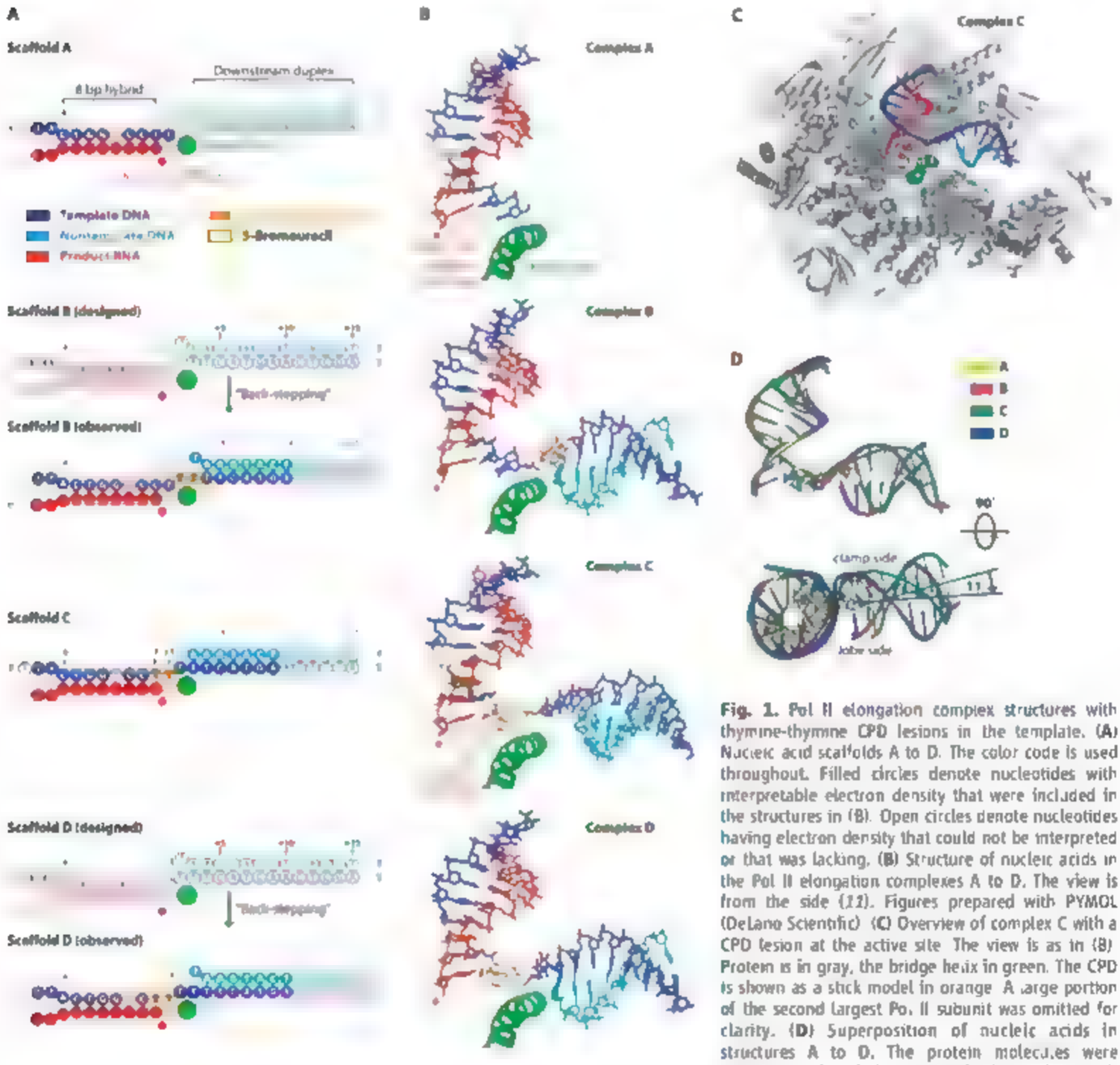


Fig. 1. Pol II elongation complex structures with thymine-thymine CPD lesions in the template. (**A**) Nucleic acid scaffolds A to D. The color code is used throughout. Filled circles denote nucleotides with interpretable electron density that were included in the structures in (B). Open circles denote nucleotides having electron density that could not be interpreted or that was lacking. (**B**) Structure of nucleic acids in the Pol II elongation complexes A to D. The view is from the side (*11*). Figures prepared with PYMOL (DeLano Scientific). (**C**) Overview of complex C with a CPD lesion at the active site. The view is as in (B). Protein is in gray, the bridge helix in green. The CPD is shown as a stick model in orange. A large portion of the second largest Pol II subunit was omitted for clarity. (**D**) Superposition of nucleic acids in structures A to D. The protein molecules were superimposed and then omitted. The nucleic acids are depicted as ribbon models, the CPDs as stick models. Upper and lower views are related by a 90° rotation around a horizontal axis.

of the scaffold, and solved the crystal structure of the resulting complex B (Fig. 1, A and B). The CPD was observed at register +2/+3, which indicated that it is not stably accommodated at positions +1/+2. Pol II had apparently "backstepped" by one position, consistent with disfavored forward translocation (Fig. 3 and SOM text). To efficiently overcome this translocation barrier, a concentration of 1 mM NTPs was required. Lower substrate concentrations limited RNA extension to one nucleotide after 5 min (fig. S2A).

To investigate the very slow incorporation of the third nucleotide into complex A, we included a CPD at positions −1/+1 and solved the structure of the resulting complex C. The CPD was seen stably accommodated in the active site, and the NTP-binding site opposite the CPD 5′-thymine was free (Fig. 1, A to C). We therefore used complex C to monitor incorporation of different NTPs. Only uridine 5′-triphosphate (UTP) led to nucleotide incorporation opposite the 5′-thymine (Fig. 2C), generally consistent with data for human Pol II (*9*). This misincorporation was very slow, with a rate constant of ~2.9 per hour, comparable with the rate determined for the third nucleotide incorporation into complex A (fig. S2B and SOM text). Because translocation is not required for nucleotide incorporation in complex C, the rate-limiting step in reaching the stalled state is the slow uridine misincorporation, not CPD translocation from positions +1/+2 to −1/+1.

Specific uridine misincorporation may be explained with the complex C structure. Whereas the CPD 3′-thymine occupies the same position as in the undamaged elongation complex (*11*), the CPD 5′-thymine is tilted by ~40° and is shifted downward by more than 2 Å into a wobble position, with the O4 atom at the location normally occupied by the N3 atom (Fig. 2D). Provided that binding of the incoming NTP (*11*, *14*) is unaffected, the wobbled 5′-thymine could form two hydrogen bonds with UTP, whereas only one hydrogen bond would be possible with other NTPs (Fig. 2D). Attempts to visualize the CPD 5′-thymine–uridine mismatch crystallographically were unsuccessful (SOM text).

These results suggested that Pol II stalls because translocation of the CPD 5′-thymine–uridine mismatch from position +1 to position −1 is strongly disfavored. This translocation event would move the damage-containing mismatch

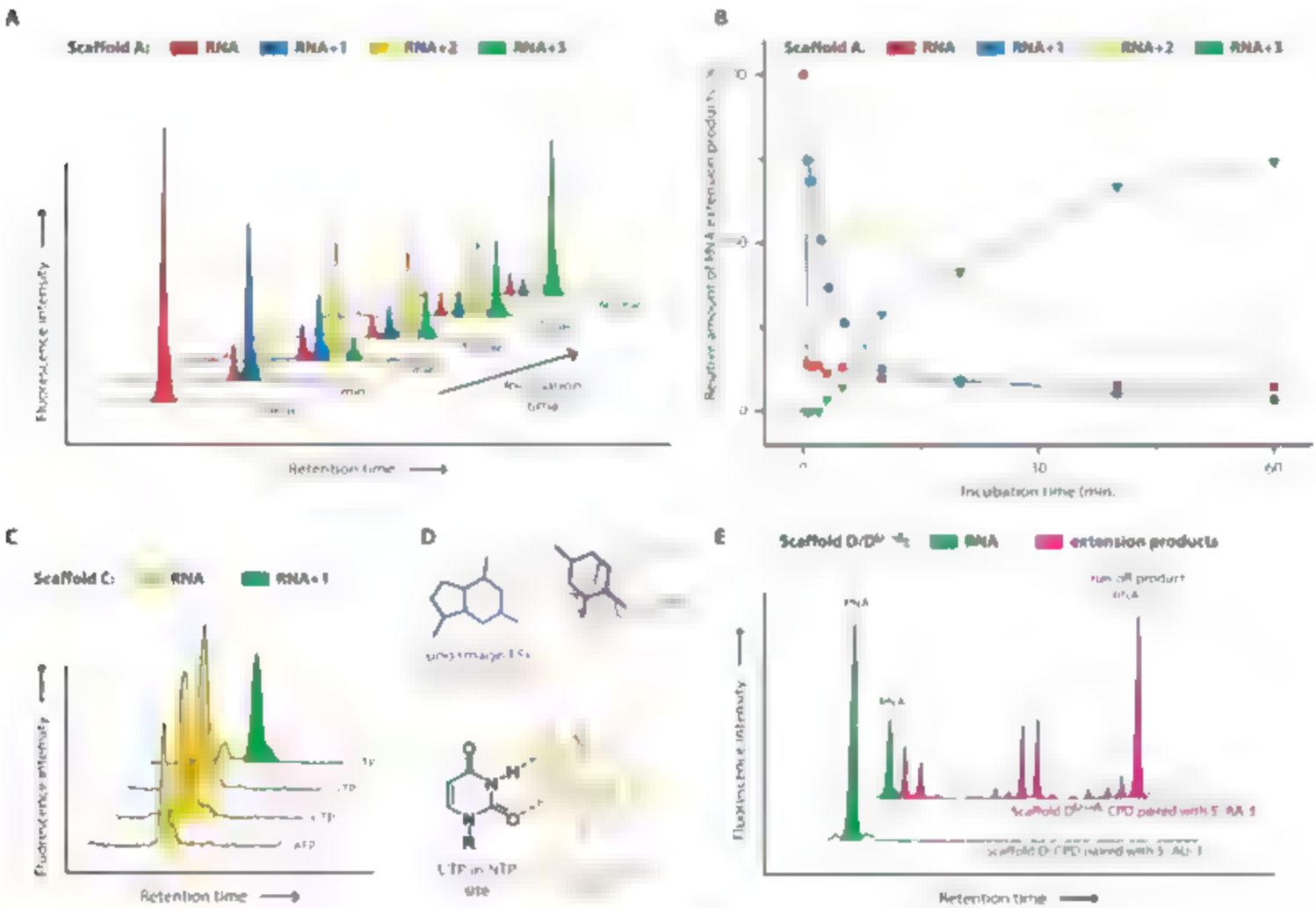


Fig. 2. RNA extension assays. (**A**) Electropherograms of time-dependent extension of RNA in complex A. A stoichiometric complex of complete Pol II and scaffold A (Fig. 1A) was incubated with 1 mM ATP, CTP, GTP, and UTP. Reactions were stopped at the given time points. RNA products were subjected to fluorescence-monitored capillary electrophoresis and identified by mass spectrometry. Signals for different RNAs are highlighted in different colors. (**B**) Quantification of time-dependent extension of RNA in complex A. Electropherogram signals in (A) were integrated, and the relative amount of RNA product was plotted against incubation time. (**C**) Specific uridine misincorporation opposite the CPD 5′-thymine. Stoichiometric complexes of complete Pol II and scaffold C were incubated with 1 mM of each NTP for 40 min. (**D**) Model for uridine misincorporation. In the upper panel, the structure of an undamaged Pol II elongation complex with a nonhydrolyzable NTP analog (PDB 1Y77) was superposed on structure C. Depicted are the base pair at position +1 in 1Y77 (violet), and the CPD in structure C (orange). As modeled in the lower panel, the CPD 5′-thymine could form two hydrogen bonds with an incoming UTP. (**E**) Lesion bypass transcription. RNA in complex D (5′-AU-3′ opposite the CPD) was not extended after a 20-min incubation with 1 mM of NTPs. Bypass was enabled under identical conditions by replacement of the RNA 3′-terminal uridine with adenosine (5′-AA-3′ opposite the CPD, scaffold D^{M-A}) (fig. S3).

into the DNA-RNA hybrid, and the resulting distortion of the hybrid would destabilize the elongation complex (*15*) (SOM text). To test this model, we incorporated the CPD at positions −2 to +1 of the scaffold, including a uridine residue opposite the 5′-thymine, and solved the structure of the resulting complex D (Fig. 1, A and B). In this structure, Pol II had apparently back-stepped by one position, and the CPD was again located at positions −1 to +1 in the active site, consistent with catalytically arrested translocation (SOM text).

Displacement translocation of the CPD from position −1 to +1 may result from distortions due to the CPD and/or due to the mismatch. To distinguish these possibilities, we tested if Pol II extends the RNA in a variant of complex D with a matched CPD 5′-thymine-adenine base pair (scaffold D^{U-A}, fig. S3). Surprisingly, this RNA was extended to the run-off transcript (Fig. 2E and SOM text). Thus, Pol II would bypass a CPD lesion if it could incorporate adenine opposite the CPD 5′-thymine. To test whether a T-U mismatch base pair alone is sufficient to stall Pol II, we used complex D, but without CPD, in RNA extension assays (scaffold $D^{\Delta CPD}$, fig. S3). Only a small portion of the RNA was extended (fig. S2D). Taken together, Pol II stalling does not result from CPD-induced distortions per se, but from CPD-directed misincorporation. In contrast, DNA polymerases can correctly incorporate adenine opposite both CPD thymines, and, depending on the type of polymerase, this can lead to stalling or lesion bypass (*16*, *17*).

In all CPD-containing structures, the polymerase conformation is unchanged. This argues against allosteric models of TCR, which assume an incoming lesion causes a conformational change in Pol II that triggers recruitment of repair factors. In complexes B and D, downstream DNA is repositioned in the polymerase cleft (Fig. 1D). However, DNA repositioning cannot support an allosteric mechanism, because it occurs only in back-stepped complexes, which would not form when NTPs are present. A damage-stalled complex could alternatively be detected via exposure of the lesion by Pol II backtracking (*18*). The transcript cleavage factor TFIIS induces backtracking of a CPD-stalled complex (fig. S2F and SOM text), but TFIIS is not required for TCR in vivo (*19*). The lesion could also be exposed after polymerase bypass or dissociation from DNA. The latter mechanism underlies bacterial TCR, which involves the adenosine triphosphatase (ATPase) Mfd (*20*). However, the related eukaryotic ATPase CSB does not trigger polymerase dissociation or bypass (*21*).

An alternative model for eukaryotic TCR that combines and extends previous models (*7*, *22*, *23*) can explain recognition of the stalled complex without allostery or exposure of the lesion (fig. S4). Complexes that stall at an arrest site are rescued by TFIIS (*24*). Complexes that stall at a nonbulky lesion are rescued by CSB-induced lesion bypass (*25*). In both cases, transcription resumes. At a CPD lesion, however, CSB counteracts TFIIS-induced backtracking (*26*, *27*), resulting in a stably stalled complex and opening a time window for assembly of the repair machinery. TFIIH catalyzes extension of the transcription bubble (SOM text). This permits dual incision of the template strand on the Pol II surface (*6*, *7*, *22*). The lesion-containing DNA fragment and the RNA transcript are removed together with Pol II, although this requires more than dual incision (*6 A*, *28*). The remaining gapped DNA is repaired. Pol II may be recycled, circumventing its ubiquitination and destruction (*29*). In conclusion, our data establish the molecular mechanism of CPD recognition by a cellular RNA polymerase, and provide a structural framework for further analysis of eukaryotic TCR.

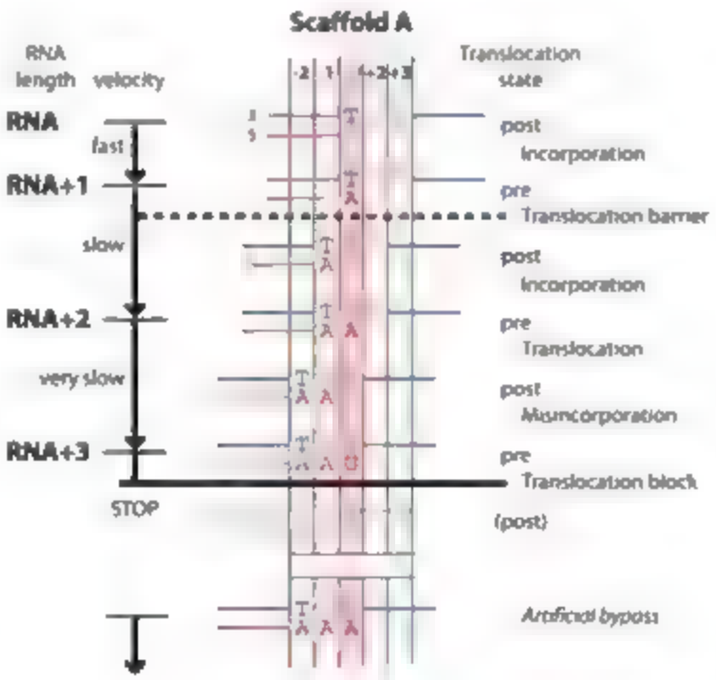


Fig. 3. Mechanism of CPD recognition by transcribing Pol II. Schematic representation of RNA extension in complex A. The initial RNA (top) corresponds to the nonextended RNA of scaffold A. The translocation barrier and the translocation block are indicated with a dashed and a solid horizontal line, respectively. The artificial situation leading to lesion bypass (Fig. 2E) is depicted at the bottom.

References and Notes

1. J. R. Mitchell, J. H. Hoeijmakers, L. J. Niedernhofer, *Curr. Opin. Cell Biol.* **15**, 232 (2003).
2. I. Mellon, G. Spivak, P. C. Hanawalt, *Cell* **51**, 241 (1987).
3. S. Tornaletti, B. A. Donahue, D. Reines, P. C. Hanawalt, *J. Biol. Chem.* **272**, 31719 (1997).
4. J. S. Me Kwen *et al. Biochem. Biophys. Res. Commun.* **320**, 1133 (2004).
5. T. T. Saxowsky, P. W. Doetsch, *Chem. Rev.* **106**, 474 (2006).
6. C. P. Selby, N. Drapkin, D. Reinberg, A. Sancar, *Nucleic Acids Res.* **25**, 787 (1997).
7. A. Tremeau-Bravard, T. Riedl, J. M. Egly, M. E. Dahmus, *J. Biol. Chem.* **279**, 7751 (2004).
8. D. Mu, A. Sancar, *J. Biol. Chem.* **272**, 7570 (1997).
9. A. Sancar, *Annu. Rev. Biochem.* **65**, 43 (1996).
10. S. Prakash, L. Prakash, *Mutat. Res.* **451**, 13 (2000).
11. H. Kettenberger, K.-J. Armache, P. Cramer, *Mol. Cell* **16**, 955 (2004).
12. Materials and methods are available as supporting material on *Science* Online.
13. A. L. Gnatt, P. Cramer, J. Fu, D. A. Bushnell, R. D. Kornberg, *Science* **292**, 1876 (2001).
14. K. D. Westover, D. A. Bushnell, R. D. Kornberg, *Cell* **119**, 481 (2004).
15. M. S. Kireeva, N. Komissarova, D. S. Waugh, M. Kashlev, *J. Biol. Chem.* **275**, 6530 (2000).
16. H. Ling, F. Boudsocq, B. S. Plosky, R. Woodgate, W. Yang, *Nature* **424**, 1083 (2003).
17. Y. Li *et al. Nat. Struct. Mol. Biol.* **11**, 784 (2004).
18. B. A. Donahue, S. Yin, J. S. Taylor, D. Reines, P. C. Hanawalt, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8502 (1994).
19. R. A. Verhage, J. Heyn, P. van de Putte, J. Brouwer, *Mol. Gen. Genet.* **254**, 284 (1997).
20. A. M. Deaconescu *et al. Cell* **124**, 507 (2006).
21. C. P. Selby, A. Sancar, *J. Biol. Chem.* **272**, 1885 (1997).
22. A. H. Sarker *et al. Mol. Cell* **20**, 187 (2005).
23. J. Q. Svejstrup, *J. Cell Sci.* **116**, 447 (2003).
24. M. Wind, D. Reines, *Bioessays* **22**, 327 (2000).
25. S. K. Lee, S. L. Yu, L. Prakash, S. Prakash, *Mol. Cell. Biol.* **22**, 4383 (2002).
26. J. P. Laine, J. M. Egly, *EMBO J.* **25**, 387 (2006).
27. C. P. Selby, A. Sancar, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11205 (1997).
28. A. Aboussekhra *et al. Cell* **80**, 859 (1995).
29. E. C. Woudstra *et al. Nature* **415**, 929 (2002).

30. We thank H. Kettenberger, K. Armache, current members of the Cramer laboratory, A. Muschielok, J. Michaelis, and C. Schulze-Briese for help. This work was supported by the Deutsche Forschungsgemeinschaft, the Sonderforschungsbereich 646, the Volkswagen-Stiftung, and the Fonds der chemischen Industrie. Part of this work was performed at the Swiss Light Source at the Paul Scherrer Institute, Villigen, Switzerland. Structure coordinates and reflection files are deposited in the Protein Data Bank under accession numbers 2ja5, 2ja6, 2ja7, and 2ja8 for complexes A, B, C, and D, respectively.

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5813/859/DC1
Materials and Methods
SOM Text
Figs. S1 to S5
Table S1
References

21 September 2006; accepted 26 December 2006
10.1126/science.1135400

Antennal Mechanosensors Mediate Flight Control in Moths

Sanjay P. Sane,[1]* Alexandre Dieudonné,[1] Mark A. Willis,[2] Thomas L. Daniel[1]

Flying insects have evolved sophisticated sensory capabilities to achieve rapid course control during aerial maneuvers. Among two-winged insects such as houseflies and their relatives, the hind wings are modified into club-shaped, mechanosensory halteres which detect Coriolis forces and thereby mediate flight stability during maneuvers. Here, we show that mechanosensory input from the antennae serves a similar role during flight in hawk moths, which are four-winged insects. The antennae of flying moths vibrate and experience Coriolis forces during aerial maneuvers. The antennal vibrations are transduced by individual units of Johnston's organs at the base of their antennae in a frequency range characteristic of the Coriolis input. Reduction of the mechanical input to Johnston's organs by removing the antennal flagellum of these moths severely disrupted their flight stability, but reattachment of the flagellum restored their flight control. The antennae thus play a crucial role in maintaining flight stability of moths.

When performing complex three-dimensional maneuvers, it is necessary for flying organisms to assess their self-motion and to respond to unintended perturbations with a rapid modulation of their wing motion (*1*, *2*). Among winged insects, diverse strategies have evolved to resolve this problem. Insects, such as dragonflies, that typically operate under brightly lit conditions may accomplish course control primarily using visual input to modulate their motor output (*3*). However, for insects flying under low light conditions, the longer latency of visual processing (*4*) means that it may not be possible for such insects to rely entirely on the visual system to assess self-motion to generate rapid maneuvers.

In Diptera, rapid acquisition of self-motion information is accomplished by halteres, club-shaped sensory structures that are evolutionarily derived from the hind wings and oscillate at wing beat frequency (*5*). The halteres and their underlying sensory neurons transduce information about pitch, roll, and yaw maneuvers to the central nervous system (*5*–*9*). Although a single muscle actively powers haltere motion with a fixed phase relation to the wing, the overall dynamics of the haltere motion is dictated primarily by its natural frequency (*5*). During rotational maneuvers, Coriolis (or gyroscopic) forces on the halteres cause a lateral deviation in their plane of motion (*5*, *9*). Because the planes of oscillation of the two halteres are nonorthogonal, this system can unambiguously distinguish between pitch, roll, or yaw rotations (*9*). Although the haltere system has received much attention because it addresses the problem of rapid course control in Dipteran flight, fewer studies have examined how non-Dipteran in-

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA. [2]Department of Biology, Case Western Reserve University, Cleveland, OH 44106, USA.

*To whom correspondence should be addressed. E-mail: sane@u.washington.edu

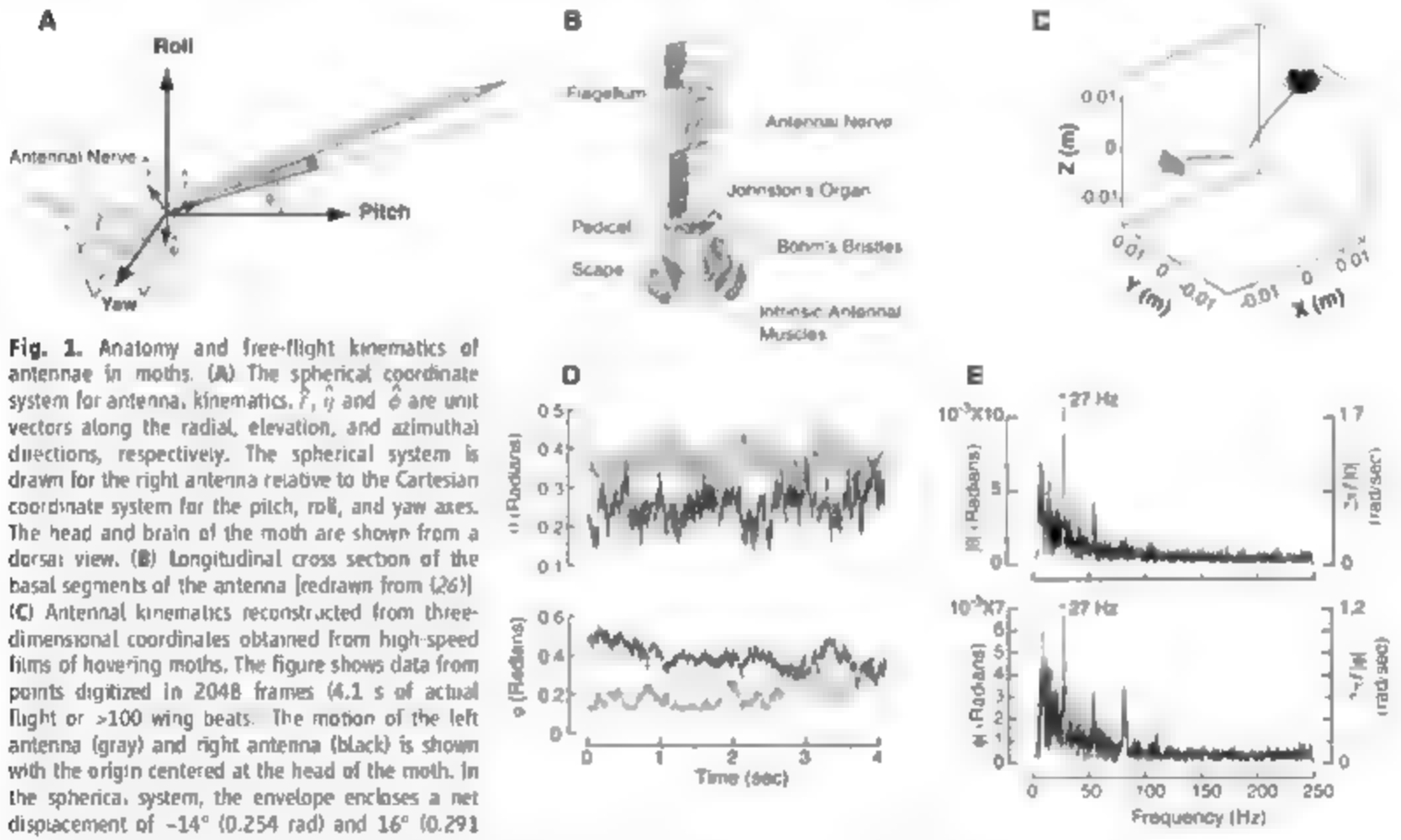


Fig. 1. Anatomy and free-flight kinematics of antennae in moths. (**A**) The spherical coordinate system for antennal kinematics. $\hat{r}$, $\hat{\theta}$ and $\hat{\phi}$ are unit vectors along the radial, elevation, and azimuthal directions, respectively. The spherical system is drawn for the right antenna relative to the Cartesian coordinate system for the pitch, roll, and yaw axes. The head and brain of the moth are shown from a dorsal view. (**B**) Longitudinal cross section of the basal segments of the antenna [redrawn from (*26*)]. (**C**) Antennal kinematics reconstructed from three-dimensional coordinates obtained from high-speed films of hovering moths. The figure shows data from points digitized in 2048 frames (4.1 s of actual flight or >100 wing beats). The motion of the left antenna (gray) and right antenna (black) is shown with the origin centered at the head of the moth. In the spherical system, the envelope encloses a net displacement of −14° (0.254 rad) and 16° (0.291 rad) along the θ angle, and −19° (0.331 rad) and 22° (0.379 rad) along the ϕ angle, for the left and right antenna, respectively. (**D**) Antennal motions along the orthogonal θ and ϕ angles. For the left antenna (gray), the pitch axis points in the opposite direction. Hence, the value of ϕ is shown after subtracting π from its actual value relative to the right-handed pitch-roll-yaw coordinate system. (**E**) Fourier decomposition of the mean offset removed movement of the right antenna for the θ, ϕ angles. The sharp peaks indicate that antennae vibrate at a wing beat frequency of 27 Hz (asterisk). The amplitude of the vibration was conservatively estimated to be 1.4° and 0.8° along the θ and ϕ angles, respectively.

sects (*10*, *11*), especially those flying under the poorly lit conditions that make visual stabilization difficult, might achieve similar maneuvering feats.

Our investigations suggest that in the crepuscular hawk moth *Manduca sexta*, which is four-winged and thus lacks halteres, the antennal mechanosensory system mediates flight stability using a mechanism similar to the halteres in Diptera. All winged insects, including moths and butterflies, possess a pair of annulated antennae (Fig. 1, A and B). Two sets of antennal mechanosensors are present on the basal antennal segments called the scape and pedicel (Fig. 1B). One set, called the Böhm's bristles, is present as fields of sensory hairs on the scape and the pedicel. In hawk moths, these fields are roughly opposite each other on each segment and roughly orthogonal to each other between two segments. The second set, called the Johnston's organ, is composed of circumferentially arranged mechanosensory stretch receptors called scolopidia. Each scolopidium is innervated by a bipolar neuron (*12*), which sends its axon down the length of the antennal nerve into the brain of the moth. The head-scape and scape-pedicel joints each consist of extrinsic (not shown) and intrinsic muscles, respectively (Fig. 1B), which drive all active antennal movements.

The antennal anatomy of *M. sexta* is common to most free-living insects (*12*). Data from silk moths (*13*) and butterflies (*14*) suggest that the Böhm's bristles encode the gross changes in antennal position, whereas the Johnston's organ responds to small, high-frequency motions of the antenna (*15*) such as vibrations due to sound or air flow (*12*, *15*, *16*). Several reports have described the possible mechanosensory function of antennae as airflow sensors (*14*, *17*–*22*). We measured the antennal motion of freely hovering *M. sexta* using high-speed videography (500 frames per second; Fig. 1, C and D) [(*23*), section S1] and found that like some other Lepidopterans [e.g., the tortoiseshell butterfly, *Aglais urticae* (*22*)], they hold their antennae at relatively constant interantennal angles (Fig. 1C) [(*23*), section S1]. About this fixed position, the antennae undergo small-amplitude vibrations (Fig. 1D) with relatively little change in the antennal length [(*23*), section S1]. The motions of the left (gray) and right (black) antenna are roughly in phase, suggesting symmetry of movement of the two antennae. Fourier transforms of these vibrations revealed prominent peaks at wing beat frequency (27 Hz; Fig. 1E). Thus, the net mechanical stimulus experienced by the antenna contains a periodic self-generated component.

The cross product of the angular velocity of body rotations and linear velocity of the moth's antennal vibrations results in Coriolis torques on the antenna. For hovering and near-hovering conditions, estimates of the torques due to Coriolis and aerodynamic forces are of the same orders of magnitude, and thus Coriolis forces also represent an important component of the net mechanical stimulus to the antennae [(*23*), section S2]. Like the halteres in Diptera (*5*, *9*), Coriolis torques thus cause lateral deflections of the antennae at either wing beat frequency

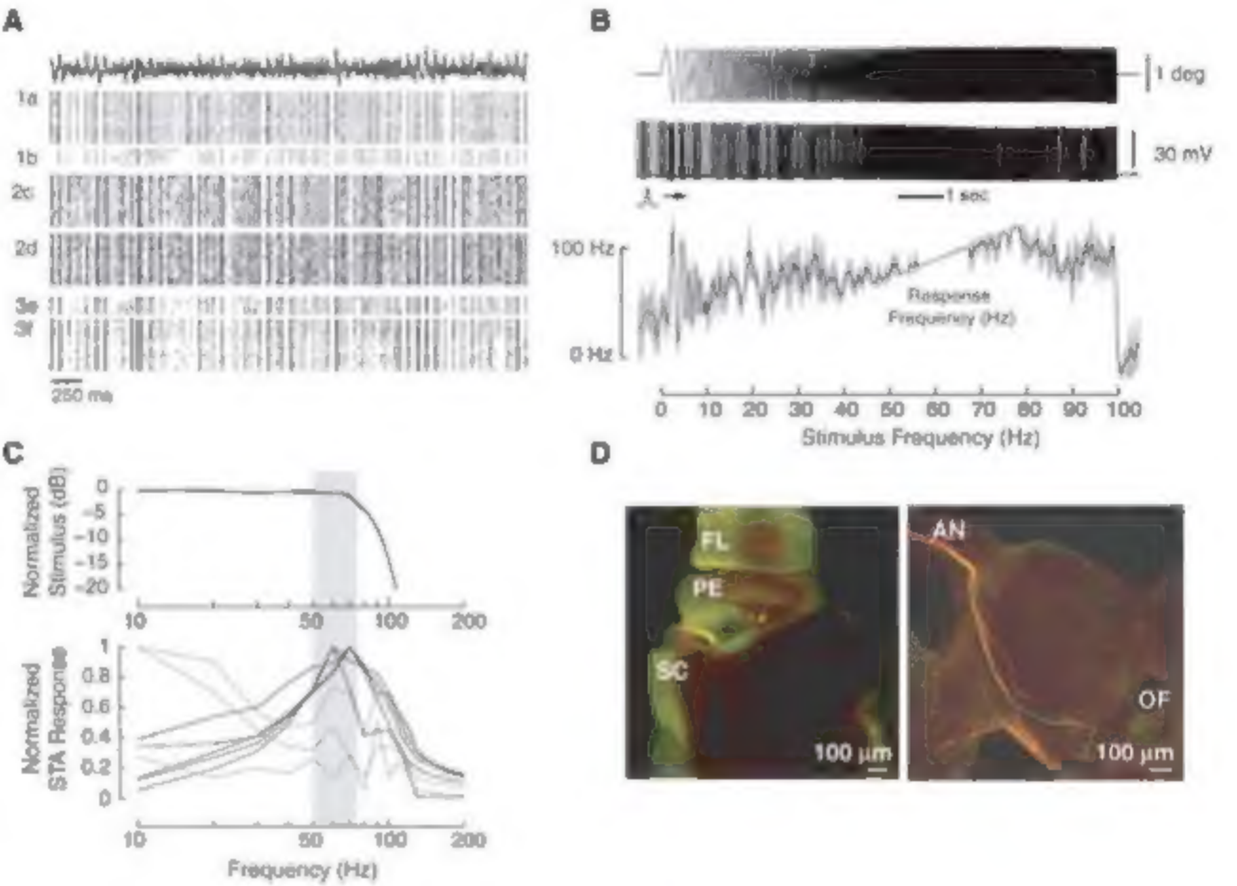


Fig. 2. Neurophysiology of mechanosensory neurons. (**A**) Spiking response of antennal mechanosensors. Raster plots of the spiking response of scolopidial neurons from the Johnston's organ in response to the "natural" stimulus in Fig. 1C. Peak antennal displacement ranged from 1° to 2.5° (three moths, six cells). The numbers in the side bar represent individual moths and letters represent individual cells. The trials consisted of repetitions of the 4-s-long "natural" stimulus repeated 4 or 12 times. (**B**) Response of an antennal scolopidium to a constant-amplitude frequency sweep from 0 to 100 Hz. A mechanical lever attached to the flagellum delivered a constant-amplitude (~1.5° peak-to-peak) frequency sweep (0 to 100 Hz, top panel) motion to the pedicel-flagellar joint. The corresponding spike response (middle panel) was measured from the axon of a scolopidial neuron and is shown as the Gaussian-convolved firing rate (black, bottom panel). The gray band enveloping this curve shows the standard deviation of the Gaussian-convolved firing rate. The width of the narrow Gaussian window is 500 ms with a standard deviation of 31.6 ms. (**C**) Stimulus-response curves for scolopidial neurons. Shown are the BLGN stimulus (top panel) and Fourier transforms for the mean spike-triggered average (STA) stimulus for eight scolopidial cells from three moths (bottom panel). The input stimulus is smoothed by the use of Welch's method and normalized with respect to maximum amplitude. The response displayed as a Fourier transform of spike-triggered average for each neuron is also normalized as a ratio of its maximum response. Of the eight neurons shown here, five (black lines) show sharp peaks between ~50 and 75 Hz and three (gray lines) show greater responses at lower frequency. (**D**) Anatomy of the scolopidial neurons. Scolopidia were filled with Alexa Fluor 568 (red) or Lucifer Yellow (yellow) dyes (left panel, four neurons). The soma of these bipolar neurons are situated in the joint between flagellum (FL) and pedicel (PE). SC denotes scape. The recordings shown in (B) were taken from the one of the neurons filled with Lucifer Yellow. Length of the scale bar is 100 μm. The neurons send their projections ipsilaterally via the antennal nerve (AN) into the antennal mechanosensory and motor center (AMMC) in the deutocerebrum of the moth, on either side of the esophageal foramen (OF) (right panel).

or twice wing beat frequency during maneuvers [(*23*), section S3; figs. S2 to S4]. Are the mechanosensors in basal segments of the antennae sufficiently sensitive for detecting the Coriolis forces on a vibrating antenna? To answer this question, we recorded intracellularly from the axons of individual scolopidial neurons in the Johnston's organ while subjecting the flagellum to controlled vibrations using a mechanical lever system [(*23*), section S4; fig. S5]. We observed strong stimulus-correlated activity of scolopidial neurons in response to mechanical stimuli in the amplitude and frequency range of natural motion (Fig. 2A). The alignment of the spikes in these records over multiple trials shows that scolopidial neurons are sensitive to small-amplitude (~1°) angular deflections in the flagella-pedicel joint. In response to constant-amplitude frequency sweep (0 to 100 Hz) motion of the antenna, the scolopidial neurons tightly phase-locked with the stimulus in a narrow frequency range of ~50 to 70 Hz. In this range, their firing rate increases linearly with the stimulus frequency (Fig. 2B). The 50- to 70-Hz window also corresponds to twice the typical range of wing beat frequency in *M. sexta* (20 to 30 Hz) (*24*) and is the expected frequency range for the Coriolis input [(*23*), section S3].

To rigorously characterize the frequency-tuning properties of the underlying sensory apparatus, we vibrated the flagellum using a band limited Gaussian noise (BLGN) waveform in a frequency window of ~0 to 100 Hz with a ~3-dB roll-off at ~85 Hz (Fig. 2C, top panel). We measured the intracellular spike responses of the scolopidial neurons and determined the pre-spike stimulus within a 100-ms time window before the occurrence of each spike of the time record. A Fourier transform of this spike-triggered average (STA) was used to estimate the stimulus frequency most likely to generate action potentials in the scolopidial neurons (*25*). A subset of the recorded scolopidial neurons (five out of eight; Fig. 2C, bottom panel) showed sharp response peaks in the frequency range of ~50 to 75 Hz and is thus capable of encoding Coriolis-driven lateral motions of the antennae during aerial maneuvers. To determine the location and projection patterns of these neurons, we filled them with fluorescent dyes after recording. In all cases, the images showed bipolar neurons innervating the scolopidia in the Johnston's organ and arborized ipsilaterally into the antennal mechanosensory and motor center (AMMC) in the brain, the main site of integration of the mechanosensory signal from the antennal deflections (Fig. 2D) (*26*, *27*).

Thus, our results suggest that the antennae of *M. sexta* (and also perhaps other Lepidopterans) are capable of detecting the Coriolis forces generated during turning maneuvers and may be involved in flight stabilization. Does a reduction of the mechanosensory input from the antenna affect flight performance? To answer this question, we performed experiments manipulating the mechanical loading on the Johnston's organs and measured the free-flight performance of the moths via videography under low-light conditions [(*23*), section S5]. This study consisted of two experiments. First, we compared the flight performance of moths with intact antennae (normally loaded Johnston's organs) to that of moths with their flagella surgically removed (unloaded Johnston's organs). Second, to control other sensory influences of an intact flagellum, we compared the performance of individuals whose flagella were removed and reattached (reloaded Johnston's organs) to that of the same individuals after removing the flagella again (unloaded Johnston's organs). In all cases, the basal mechanosensory apparatus was left intact and only the flagellum was cut within a few (<10) annuli from the base. This ensured that the mechanical stimulus due to vibrating antennae was substantially reduced.

To analyze the effect of the flagellar amputation on flight performance, we counted the occurrences of three types of behavior in each flight sequence: (i) collisions with the wall, (ii) crashes to the floor, and (iii) incidence of backward flight (typically spanning the entire width of the flight chamber) [(*23*), sections S5 to S7]. Moths with amputated flagella could take flight, but had a significantly higher frequency of backward flight than intact moths [(*23*), section S7; Fig. 3A]. These moths also crashed [(*23*), section S7; Fig. 3B] and collided [(*23*), section S7; Fig. 3C] with the walls significantly more frequently than moths with intact flagella ($P < 0.02$). These results indicate that intact flagella are necessary for stable flight in *M. sexta*.

The flagella of most insect antennae are multisensory structures bearing odor, humidity, and temperature sensors in addition to mechanosensory bristles along the flagellum (*12*). Thus, in moths with amputated flagella, not only is the normal mechanical loading of Johnston's organs disrupted, but these various sensory cues are also eliminated. Which of these functions is primarily responsible for the lack of flight stability in moths with amputated flagella? To answer this question, we reattached the flagella to the basal antennal segments of the moths with glue and compared their flight performance with moths without flagella [(*23*), section S5]. Moths with reattached flagella were able to substantially regain flight control [(*23*), section S8]. In contrast, moths with no flagella flew backward significantly more often (Fig. 3D), crashed to the ground more often (Fig. 3E), and collided with walls more frequently (Fig. 3F). Because the surgical removal and reattachment of the antennal flagella with Super Glue required cold anesthesia, these moths were analyzed separately from the earlier flagella-amputated group. Moths

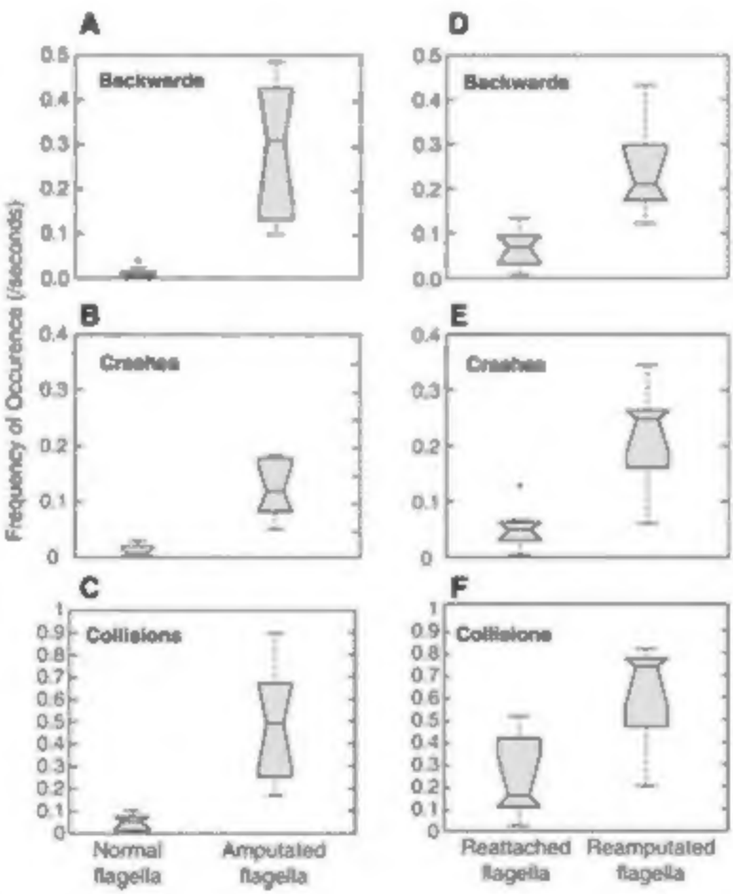


Fig. 3. Free-flight behavior of moths with and without antennal mechanosensory input. Data from behavioral experiments are represented as notched, whisker plots. The bottom and top of the box show lower and upper quartile values, respectively. The horizontal black line within the box represents the median for each category. Whiskers include the most extreme data value up to 1.5 times the interquartile range. The outlier points beyond this range are shown by the + symbol. Nonoverlapping notches indicate different medians at the 5% significance level. Top row shows data for backward flight, middle row for crashes, and bottom row for collisions. The left column (**A** to **C**) shows behavior of moths with intact antennae versus amputated flagella (Wilcoxon signed rank test, $P = 0.002$). The right column (**D** to **F**) shows behavior of moths with reattached flagella versus reamputated flagella (Wilcoxon signed rank test, $P = 0.0156$).

with reattached flagella receive no input from any sensory neurons on the flagella because their antennal nerve is severed during the removal and reattachment of flagella. Thus, the restoration of normal mechanical loading of Johnston's organs is sufficient for recovery of flight control in moths with reattached flagella.

In summary, our study of antennal vibrations in freely flying moths and subsequent estimates of the mechanical forces on vibrating antennae suggest that antennae experience strong Coriolis forces during aerial maneuvers. Our neurophysiological data indicate that antennal mechanosensors are tuned within the frequency range of these forces. The antennectomy experiments show that normal loading of the mechanosensors in the base of the antennae, probably the Johnston's organ, is critical for flight stability. Moths without flagella can take flight but cannot stably hover or execute controlled maneuvers. However, moths with reattached flagella regain flight control. Therefore, we propose that the antennae together with the mechanosensors in their bases are necessary for flight stability in *M. sexta*, similar to halteres in flies. These data provide a mechanism to explain previous observations that antennectomy compromises flight in other Lepidopterans (*21*). Moreover, it is possible that the end knobs on the antennae of certain Lepidoptera (e.g., butterflies) and other insects (e.g., owl flies) function as a means of increasing the sensitivity to Coriolis forces by enhancing antennal vibrations during flight. Thus, these studies offer an insight into the non haltere-mediated mechanisms for course control in flying insects and may be useful to studies of insect flight, as well as to recent efforts toward designing robotic insects.

References and Sections

1. T. S. Collett, M. F. Land, *J. Comp. Physiol.* **99**, 1 (1975).
2. H. Wagner, *Philos. Trans. R. Soc. London B Biol. Sci.* **312**, 553 (1986).
3. R. M. Olberg, *J. Comp. Physiol.* **141**, 327 (1981).
4. E. J. Warrant, *Vision Res.* **39**, 1611 (1999).
5. J. Pringle, *Philos. Trans. R. Soc. London B Biol. Sci.* **233**, 347 (1948).
6. G. Nalbach, R. Hengstenberg, *J. Comp. Physiol.* **175**, 695 (1994).
7. W. P. Chan, F. Prete, M. H. Dickinson, *Science* **280**, 289 (1998).
8. M. H. Dickinson, *Philos. Trans. R. Soc. London B Biol. Sci.* **354**, 903 (1999).
9. G. Nalbach, *J. Comp. Physiol.* **173**, 293 (1993).
10. W. Pfa, G. Nalbach, J. Zeil, *Naturwissenschaften* **80**, 371 (1993).
11. V. B. Wigglesworth, *Nature* **157**, 655 (1946).
12. D. Schneider, *Annu. Rev. Entomol.* **9**, 103 (1964).
13. D. Schneider, K. E. Kaissling, *Zool. Jahrb. Abt. Anat. Ontog. Tiere* **75**, 287 (1956).
14. M. Niehaus, M. Gewecke, *Zoomorphologie* **91**, 19 (1978).
15. L. H. Field, T. Matheson, *Adv. Insect Physiol.* **27**, 86 (1998).
16. D. Robert, M. C. Goplen, *J. Insect Physiol.* **48**, 189 (2002).
17. M. Gewecke, in *Experimental Analysis of Insect Behaviour*, L. B. Browne, Ed. (Springer, Berlin, Heidelberg, New York, 1974), pp. 100–113.
18. M. Gewecke, *Nature* **225**, 1263 (1970).
19. H. G. Heinzel, M. Gewecke, *J. Comp. Physiol.* **161**, 671 (1987).
20. M. Gewecke, H. G. Heinzel, *J. Comp. Physiol.* **139**, 357 (1980).
21. M. Niehaus, *J. Comp. Physiol.* **145**, 257 (1981).
22. M. Gewecke, M. Niehaus, *J. Comp. Physiol.* **145**, 249 (1981).
23. Materials and methods are available as supporting material on *Science* Online.
24. S. P. Sane, N. P. Jacobson, *J. Exp. Biol.* **209**, 43 (2006).
25. P. Dayan, L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, MA, 2005).
26. P. Kloppenburg, S. M. Camazine, X. J. Sun, P. Randolph, J. G. Hildebrand, *Cell Tissue Res.* **287**, 425 (1997).
27. J. S. Vande Berg, *J. Morphol.* **133**, 439 (1971).

28. We gratefully acknowledge M. Tu for sharing high-speed videos of hovering moths; T. Hedrick for help with digitization; A. Hinterwirth and A. Mountcastle for assistance with behavior experiments; Sam and Suma Chilukuri for their support; and an anonymous referee for useful feedback. Support was provided by an NSF Inter-Disciplinary Informatics Grant (S.P.S.) and the Office of Naval Research (I.L.D. and M.A.W.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/315/5813/863/DC1
Materials and Methods
Figs. S1 to S5
Movies S1 to S3

8 August 2006; accepted 8 January 2007
10.1126/science.1133598

Living Cell Research Microscope

The Axio Observer has been developed for research of processes in living cells. Based on an inverted research microscope, the Axio Observer makes it possible not only to observe processes in living cells, but also to manipulate processes and analyze the resulting changes. Great emphasis has been laid on optimum operating convenience. For example, the ACR function automatically recognizes objectives and reflector modules, integrates them into the system configuration, and displays them on the touchscreen display. This feature is particularly helpful when several researchers share one instrument. The six-position reflector turret allows the fast change of filter sets in the beam path during an experiment. Open system architecture makes it easy to integrate external components. For example, the laser port enables the use of a laser for targeted manipulation in fluorescence recovery after photobleaching (FRAP) procedures or targeted deletion of cellular structures. In addition to mechanical manipulators, optical tools such as laser scan modules, optical tweezers, and laser microdissection transform the Axio Observer into a complete cell research station.
Carl Zeiss. For information 914-681-7627 www.zeiss.com

Mitochondria Isolation

The Focus Mitochondria Kit enables the fast and simple isolation of enriched mitochondrial fractions from animal cells and tissue. Most of the mitochrondria isolated using this kit contains intact inner and outer membranes.
G-Biosciences/Genotech For information 800-628-7730 www.GBiosciences.com

Laser Scanning Microscope Update

The LSM 5 DUO Release 4.2 is an enhanced version of a laser scanning microscope (LSM) for examinations in biomedical research. This new version makes work with many new fluorescent dyes faster and more efficient. The benefits of the LSM 5 DUO systems—flexibility, simultaneous sample manipulation, and the versatility of a real confocal workstation—have been extended by improved laser modules. The system also provides an increased resolution of 2.4 megapixels in the fast line scan mode and an integrated autofocus function. The LSM 5 DUO is based on the inverted microscope AxioObserver.Z1, which not only features a fully integrated incubation system, but also allows use of a wide variety of reflector modules. The complete reflector turrets with the DUO beam combiners can be easily exchanged.
Carl Zeiss For information 914-681-7627 www.zeiss.com/lsm

FT-IR Imaging Accessory

The attenuated total reflectance (ATR) image accessory for the Spotlight FT-IR (Fourier transform-infrared) imaging system is designed to provide high quality data while maximizing the image area. ATR is a surface measurement technique that extends the range of samples that can be measured on FT-IR systems, enabling laboratories to get meaningful FT-IR spectra from samples that were impossible using traditional transmission or reflectance measurements. The accessory can resolve material sample features as small as 3.1 μm. It features a highly efficient germanium crystal with a 600-μm diameter sample area that enables laboratories to measure large samples (such as multi-layer laminate samples a few hundred μm across) with excellent signal quality in one operation without having to resort to "stitching" several images.
PerkinElmer For information 781-237-5100 www.perkinelmer.com/ATR

Protein Expression Medium

Protein Expression Medium is a serum-free, animal origin-free medium developed for the efficient production of recombinant proteins in PER.C6 cells. This medium supports the adaptation, growth, and expansion of transfected PER.C6 human embryonic retinoblast cells for large-scale, high-density suspension culture in shake flasks, roller bottles, and bioreactors. It also supports parental PER.C6, HEK 293, an HeLa cells in suspension culture for protein expression applications.
Invitrogen For information 800-955-6288 www.invitrogen.com

Unstable DNA Cloning

CopyCutter EPI400 *E. coli* cells significantly lower the copy number of many common vectors so that "toxic" genes or unstable DNA sequences are easier to clone. Reducing the copy number of vectors increases the likelihood that an insert will be stably maintained.
Epicentre Biotechnologies For information 800-284-8474 www.EpiBio.com/Copycutter.asp

Mutation Surveyor Software

Mutation Surveyor is analytical software for the detection of DNA variants in sequence traces. It features a patent-pending, anti-correlation technology that provides up to 99.5% accuracy, eliminates false negatives, and limits false positives to less than 50 ppm. Version 3.01 includes two additional functionalities: vector masking for use in the quality control of DNA clones and negative single nucleotide polymorphism (SNP) indicator for use in clinical analysis of sequence traces. The vector masking tool allows users to inform the software of the vector region, which permits the software to make use of its patented physical trace comparison technology to detect any sequence change from the original sequence trace in up to 400 clone traces in less than 90 seconds. The negative SNP indicator indicates the absence of a nucleotide change in patient sequence traces.
SoftGenetics For information 814-237-9340 www.softgenetics.com

For more information visit Product-Info, *Science*'s new online product index at *http://science.labvelocity.com*

From the pages of Product-Info, you can:
- Quickly find and request free information on products and services found in the pages of *Science*.
- Ask vendors to contact you with more information.
- Link directly to vendors' Web sites.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and government organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier by visiting www.science.labvelocity.com on the Web, where you can request that the information be sent to you by e-mail, fax, mail, or telephone.